STIC Search Report
Biotech-Chem Library

To: Fozia Hamud Location: rem/4d64/4c70 Art Unit: 1647 Wednesday, July 07, 2004	From: Beverly Shears Location: Remsen Bldg. RM 1A54 Phone: 571-272-2528 beverly.shears@uspto.gov

Search Notes

Case Serial Numbers 10/076260 and 10/030226 attached.

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

CW protein - protein search, using sw model

Run on: June 30, 2004, 17:14:53 ; Search time 80 Seconds
(without alignments)
1222.018 Million cell updates/sec

Title: US-10-076-260-2
Perfect score: 1853
Sequence: 1 MYNGSCCRIEGDTISQVMPP.........ANSFQSQSDGQWDPHIVEWH 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1853	100.0	346	4	AAG80968	Aag80968 Human nGP
2	1853	100.0	346	4	ABB44522	Abb44522 Human GPC
3	1853	100.0	346	4	AAJ06197	Aau06197 Novel hum
4	1853	100.0	346	4	AAJ04373	Aau04373 Human G-p
5	1853	100.0	346	5	AAE16172	Aae16172 Human G-p
6	1853	100.0	346	5	AAU11401	Aau11401 HM74-like
7	1853	100.0	346	5	AAE17077	Aae17077 Human G-p
8	1853	100.0	346	5	ABB08596	Abb08596 Human lip
9	1853	100.0	346	5	ABG93786	Abg93786 Human G p
10	1853	100.0	346	5	ABP95599	Abp95599 Human GPC
11	1853	100.0	346	5	AAO14788	Aao14788 Human pur
12	1853	100.0	346	5	AAE24354	Aae24354 Human G p
13	1853	100.0	346	6	ABP81747	Abp81747 Human che
14	1853	100.0	346	6	ABP56751	Abp56751 Human GAV
15	1853	100.0	346	6	AAO26511	Aao26511 Human G-P
16	1853	100.0	346	6	ABP58453	Abp58453 Human res
17	1853	100.0	346	7	ADC46872	Adc46872 Human TA-
18	1853	100.0	346	7	ABW00810	Abw00810 Human GPC
19	1853	100.0	346	7	ADE40282	Ade40282 Human NOV
20	1853	100.0	346	7	ADE40272	Ade40272 Human NOV
21	1853	100.0	346	7	ADE40278	Ade40278 Human NOV
22	1853	100.0	352	7	ADE40286	Ade40286 Human NOV
23	1849	99.8	346	7	ABB82502	Abb82502 Human TGR
24	1846	99.6	346	7	ADE40274	Ade40274 Human NOV
25	1839	99.2	346	4	ABB44523	Abb44523 Human GPC
26	1839	99.2	346	7	ADE40276	Ade40276 Human NOV
27	1739	93.8	342	5	ABP54312	Abp54312 Human G p
28	1686	91.0	314	7	ADE40280	Ade40280 Human NOV
29	1686	91.0	320	7	ADE40288	Ade40288 Human NOV
30	1686	91.0	320	7	ADE40284	Ade40284 Human NOV
31	1570	84.7	296	5	AAG80934	Aag80934 Human nGP
32	1570	84.7	296	5	ABG93752	Abg93752 Human G p
33	1481	79.9	351	6	AAO26512	Aao26512 Mouse G-P
34	1477	79.7	351	6	AAO26513	Aao26513 Rat G-Pro
35	1272	68.6	247	7	ADE40290	Ade40290 Human NOV
36	1105	59.6	215	7	ADE40292	Ade40292 Human NOV
37	904	48.8	363	5	ABG72361	Abg72361 Rabbit or
38	897.5	48.4	387	5	AAU77992	Aau77992 Human inf
39	889	48.0	363	6	ABG72360	Abg72360 Monkey or
40	883.5	47.7	362	6	ABG72358	Abg72358 Human orp
41	883.5	47.7	363	2	AAW94654	Aaw94654 G-protein
42	883.5	47.7	363	4	AAU04379	Aau04379 Human G-p
43	883.5	47.7	363	4	ADC86215	Adc86215 Human GPC
44	881.5	47.6	387	3	AAY90672	Aay90672 Human mut
45	881.5	47.6	387	6	ABR48195	Abr48195 Human bla

ALIGNMENTS

RESULT 1
AAG80968
ID   AAG80968 standard; protein; 346 AA.
XX
AC   AAG80968;
XX
DT   28-AUG-2001 (first entry)
XX
DE   Human nGPCR11 #2.
XX
KW   G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW   signal transduction; schizophrenia; thyroid disorder; renal failure;
KW   rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW   cardiovascular disease; proliferative disorder; hormonal disorder;
KW   neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW   attention deficit-hyperactivity disorder/attention deficit disorder;
KW   Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW   rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW   neuroprotective.
XX
OS   Homo sapiens.
XX
PN   WO200136473-A2;
XX
PD   25-MAY-2001;
XX
PF   16-NOV-2000; 2000WO-US031581.
XX
PR   16-NOV-1999;   99US-0165838P.
PR   17-NOV-1999;   99US-0166071P.
PR   19-NOV-1999;   99US-0166678P.
PR   28-DEC-1999;   99US-0173396P.
PR   22-FEB-2000;   2000US-0184129P.
PR   28-FEB-2000;   2000US-0185421P.
PR   28-FEB-2000;   2000US-0185554P.
PR   02-MAR-2000;   2000US-0186530P.
PR   03-MAR-2000;   2000US-0186811P.
PR   09-MAR-2000;   2000US-0188114P.
PR   17-MAR-2000;   2000US-0190310P.
PR   21-MAR-2000;   2000US-0190800P.
PR   02-APR-2000;   2000US-0198568P.
PR   02-MAY-2000;   2000US-0201190P.
PR   08-MAY-2000;   2000US-0203111P.
PR   25-MAY-2000;   2000US-0207094P.
XX
PA   (PHAA) PHARMACIA & UPJOHN CO.
XX
PI   Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;

```
PI   Schellin KA, Kaytes PS, Barnigan CM, Ruff V, Sejlitz T, Huff RM;
XX
DR   WPI; 2001-389826/41.
DR   N-PSDB; AAH51008.
XX
PT   New G protein-coupled receptor (nGPCR-x) and its encoding polynucleotide
PT   useful for diagnosing and treating e.g. schizophrenia.
XX
PS   Claim 37; Page 89; 261pp; English.
XX
CC   The present invention relates to novel G protein-coupled receptors
CC   (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
CC   31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
CC   one such G protein-coupled receptor. GPCRs are also known as seven
CC   transmembrane receptors and function in signal transduction. The nGPCRx
CC   coding sequences are useful for screening a human to diagnose a disorder
CC   affecting the brain or a genetic predisposition, specifically
CC   schizophrenia. nGPCRx are useful for identifying compounds useful for
CC   treating schizophrenia. Detection of nGPCRx in a sample is useful as a
CC   diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
CC   failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
CC   metabolic and cardiovascular diseases, proliferative disorders and
CC   hormonal disorders. Modulators of nGPCRx activity have the utility for
CC   treating neurological disorders, including schizophrenia, ADHD/ADD
CC   (attention deficit-hyperactivity disorder/attention deficit disorder),
CC   and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
CC   migraine and senile dementia. Additional disorders include inflammatory
CC   conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
CC   disorders, cancers, respiratory ailments such as asthma, and inflammatory
CC   diseases e.g. inflammatory bowel disease
XX
SQ   Sequence 346 AA;

Query Match             100.0%;  Score 1853;  DB 4;  Length 346;
Best Local Similarity   100.0%;
Matches 346;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFEHMKTWKPSTVYLFNLAVA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFEHMKTWKPSTVYLFNLAVA 60

Qy    61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Qy   121 HHAVNTISTRVAAGIVCTLMALVIIGTVYLLLRENHLCVQETAVSCESFIMBSANGWHDIM 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 HHAVNTISTRVAAGIVCTLMALVIIGTVYLLLRENHLCVQETAVSCESFIMBSANGWHDIM 180

Qy   181 FQLBFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 FQLBFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240

Qy   241 LYFIWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 LYFIWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK 300

Qy   301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346

RESULT 2
ABB44522
ID   ABB44522 standard; protein; 346 AA.
XX
AC   ABB44522;
XX
DT   28-JAN-2002  (first entry)
XX
DE   Human GPCR1a polypeptide SEQ ID NO 2.
XX
KW   Human; GCPR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
```

Second column:

```
KW   anabolic; cytostatic; antiviral; gene therapy; cardiomyopathy; obesity;
KW   anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
KW   asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;
KW   infection; human immunodeficiency virus; HIV.
XX
OS   Homo sapiens.
XX
PN   WO200174904-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US010241.
XX
PR   31-MAR-2000; 2000US-0193664P.
PR   05-APR-2000; 2000US-0194614P.
PR   06-APR-2000; 2000US-0195063P.
PR   06-APR-2000; 2000US-0195066P.
PR   06-APR-2000; 2000US-0195067P.
PR   06-APR-2000; 2000US-0195068P.
PR   06-APR-2000; 2000US-0195069P.
PR   06-APR-2000; 2000US-0195070P.
PR   06-APR-2000; 2000US-0195510P.
PR   21-JUL-2000; 2000US-0219855P.
PR   27-JUL-2000; 2000US-0221284P.
PR   28-JUL-2000; 2000US-0221325P.
PR   11-AUG-2000; 2000US-0224588P.
PR   11-OCT-2000; 2000US-0239613P.
PR   18-JAN-2001; 2001US-0262508P.
PR   23-JAN-2001; 2001US-0263433P.
PR   23-JAN-2001; 2001US-0263604P.
PR   30-JAN-2001; 2001US-0265161P.
PR   29-MAR-2001; 2001US-0082317J.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Majumder K,   Vernet CAM,   Casman SJ,   Wolenc AR,   Spaderna SK,
PI   Padigaru M,   Mishmu VS,   Tchernev VT,   Spytek KA,   Li L,   Baumgartner JC;
PI   Gusev VY;
XX
DR   WPI; 2001-639351/73.
DR   N-PSDB; ABA81529, ABA81530.
XX
PT   New human G-protein coupled receptor X, GPCRX, polypeptide useful in
PT   treatment or prevention of GPCRX associated disorders e.g. cardiomyopathy
PT   or atherosclerosis, and to screen for antagonists and agonists useful
PT   therapeutically.
XX
PS   Claim 1; Page 8; 157pp; English.
XX
CC   The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC   encode G-coupled protein-receptor related polypeptides (ABB44529-
CC   ABB44543). The isolated polypeptide having a sequence differing by no
CC   more than 15 % of amino acid residues from one of 22 amino acid sequences
CC   (or mature forms of the sequences), fully defined in the specification
CC   and corresponding to human G-protein coupled receptor X (GPCRX)
CC   polypeptides. The polypeptides have potential cardiant,
CC   antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The
CC   polypeptides can be administered therapeutically, especially using gene
CC   therapy and expressing the encoding DNA in vivo, to treat or prevent
CC   GPCRX-associated disorders, especially in humans. For example, they can
CC   be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC   related to signal processing and metabolic pathway modulation (e.g.
CC   obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
CC   sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC   disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC   haematopoietic disorders, developmental diseases, neurological disorders,
CC   bacterial, fungal, protozoal and viral infections (e.g. with human
CC   immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
CC   to determine the presence of or predisposition to a disease associated
CC   with altered levels of the polypeptide in mammals (especially humans) by
CC   detecting alterations in polypeptide expression levels relative to
CC   control samples. They are useful to identify agents binding polypeptide
CC   (e.g. cellular receptors or downstream effectors) and/or agents
```

```
Db        61  DFLLMICLPFRTDYYLRRRFWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy       121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180

Qy       181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240

Qy       241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK  300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK  300

Qy       301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
```

Search completed: June 30, 2004, 17:21:53
Job time : 84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 03:46:36 ; Search time 516 Seconds
(without alignments)
8545.794 Million cell updates/sec

Title: US-10-076-260-1
Perfect score: 1038
Sequence: 1 atgtacaacggtcgtgctg........cccacattgttgagtggcac 1038

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1038	100.0	1038	6	ABA99236	Aba99236 Human G p
2	1038	100.0	1038	7	AAL53846	Aal53846 DNA of hu
3	1038	100.0	1041	4	AAH51008	Aah51008 Human nGP
4	1038	100.0	1041	5	AAS12581	Aas12581 cDNA enco
5	1038	100.0	1041	5	AAS07946	Aas07946 Human cDN
6	1038	100.0	1041	6	AAD27497	Aad27497 Human G-p
7	1038	100.0	1041	6	ABS70241	Abs70241 DNA encod
8	1038	100.0	1041	6	ABZ42873	Abz42873 Human GPC
9	1038	100.0	1041	6	AAL42499	Aal42499 Human pur
10	1038	100.0	1041	7	ABZ22648	Abz22648 Human GAV
11	1038	100.0	1041	7	ABZ57957	Abz57957 Human res
12	1038	100.0	1041	9	AAD61654	Aad61654 Human GPC
13	1038	100.0	1050	4	ABA81530	Aba81530 Human GPC
14	1038	100.0	1050	4	ABA81529	Aba81529 Human GPC
15	1038	100.0	1050	9	ADE40271	Ade40271 Human NOV
16	1038	100.0	1057	9	ADE40285	Ade40285 Human NOV
17	1038	100.0	1057	9	ADE40277	Ade40277 Human NOV
18	1038	100.0	1060	9	ADE40281	Ade40281 Human NOV
19	1038	100.0	1083	6	AAD26371	Aad26371 Human G-p
20	1038	100.0	1372	7	ABZ42592	Abz42592 Human che
21	1038	100.0	1730	6	AAS18501	Aas18501 cDNA enco
22	1038	100.0	2331	5	AAS12582	Aas12582 Gene enco
23	1038	100.0	2345	9	ADC46871	Adc46871 Human TA-
24	1038	100.0	3612	9	ADC46870	Adc46870 Human TA-
25	1036.4	99.8	1104	9	ADE40273	Ade40273 Human NOV
26	1036.4	99.8	1194	9	AAD39181	Aad39181 Human G p
27	1034.8	99.7	1041	7	ABV73367	Abv73367 Human TGR
28	1034.8	99.7	1104	4	ABA81531	Aba81531 Human GPC
29	1034.8	99.7	1104	9	ADE40275	Ade40275 Human NOV
30	1014	97.7	2580	6	ABQ81131	Abq81131 Human G p
31	942.4	90.8	961	9	ADE40283	Ade40283 Human NOV
32	942	90.8	961	9	ADE40287	Ade40287 Human NOV
33	942	90.8	961	9	ADE40279	Ade40279 Human NOV
34	888	85.5	888	4	AAH50974	Aah50974 Human nGP
35	888	85.5	888	6	ABS70207	Abs70207 DNA encod
36	724.4	69.8	1668	6	AAL53848	Aal53848 DNA of ra
37	723.4	69.7	742	9	ADE40289	Ade40289 Human NOV
38	722.8	69.6	3251	7	AAL53847	Aal53847 DNA of mo
39	627.4	60.4	646	9	ADE40291	Ade40291 Human NOV
40	627.4	60.4	646	9	ADE40297	Ade40297 Human NOV
41	402.8	38.8	1254	6	ABS57853	Abs57853 Rabbit cD
42	372.4	35.9	1174	6	ABK47759	Abk47759 cDNA enco
43	372.4	35.9	1361	2	AAX16671	Aax16671 G-protein
44	369.2	35.6	1092	5	AAS07952	Aas07952 Human CDN
45	369.2	35.6	1492	9	ADC86214	Adc86214 Human GPC

ALIGNMENTS

RESULT 1
ABA99236
ID ABA99236 standard; DNA; 1038 BP.
XX
AC ABA99236;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human G protein-coupled receptor protein TGR13 encoding sequence.
XX
KW Antiinflammatory; anorectic; obesity; inflammation; gene therapy; human;
KW G protein-coupled receptor protein TGR13; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1038
FT /*tag= a
FT /product= "G-protein-coupled receptor protein TGR13"
FT /partial
FT /note= "no stop codon"
XX
PN WO200202767-A1.
XX
PD 10-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-JP005711.
XX
PR 04-JUL-2000; 2000JP-00206860.
PR 31-JUL-2000; 2000JP-00235274.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Shibata S, Horikoshi K, Taniyama Y, Shintani Y, Miyajima N;
XX
DR WPI; 2002-164535/21.
DR P-PSDB; ABB08596.
XX
PT New human lipocyte-originated G protein-coupled receptor proteins TGR13
PT and encoding DNAs, for developing drugs to treat obesity and
PT inflammations, including gene therapy.
XX
PS Claim 6; Page 94-95; 101pp; Japanese.
XX
CC This invention relates to a human lipocyte-originated G protein-coupled
CC receptor proteins TGR13, thought to be antiinflammatory and anorectic in

```
CC  their action. The proteins and encoded DNAs are for use in developing
CC  drugs to treat obesity and inflammation, including gene therapy. The
CC  present sequence represents the human lipocyte-originated G protein-
CC  coupled receptor protein TGR13 encoding sequence
XX
SQ  Sequence 1038 BP; 207 A; 294 C; 277 G; 260 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1038; DB 6;  Length 1038;
Best Local Similarity   100.0%;  Pred. No. 5e-287;
Matches 1038;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGTACAACGGGTCGTGCTGCCGCCATCGAGGGGGACCATCTCCCAGGTGATGCCGCCG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGTACAACGGGTCGTGCTGCCGCCATCGAGGGGGACCATCTCCCAGGTGATGCCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCGCAATGGGGTCGCCCGTGTGGTTTC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCGCAATGGGGTCGCCCGTGTGGTTTC 120

Qy  121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTACCTTTTCAATTTGGCCGTGGCT 180

Qy  181 GATTTCCTCCTTATGATCTGCCTGCCCTTTCGGACAGCATTACCTCAGACGTAGACAC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GATTTCCTCCTTATGATCTGCCTGCCCTTTCGGACAGCATTACCTCAGACGTAGACAC 240

Qy  241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGGC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGGC 300

Qy  301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCACCCC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCACCCC 360

Qy  361 CACCACGCGGTGAACACTATCTCCACCCGGTGGCGGCTGGCATGTCTGCACCCTGTGG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CACCACGCGGTGAACACTATCTCCACCCGGTGGCGGCTGGCATGTCTGCACCCTGTGG 420

Qy  421 GCCCTGGTCATCCTGGGACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCCTGGTCATCCTGGGACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG 480

Qy  481 ACGGCCGTCCTCCTGGAGAGCTTCATCATGTGGCATGACATCATG 540
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ACGGCCGTCCTCCTGGAGAGCTTCATCATGTGGCATGACATCATG 540

Qy  541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTTATTTTGCTCCTTCAAGATTGTT 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTTATTTTGCTCCTTCAAGATTGTT 600

Qy  601 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGGACGCTGGCCGGATGAAGAAGGCGACCCG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGGACGCTGGCCGGATGAAGAAGGCGACCCG 660

Qy  661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy  721 CTCTATTTCCTCTGAGCCGTGCCCTGCGAGTGCCTGCATCCCTGTCATGGGGCCCTG 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CTCTATTTCCTCTGAGCCGTGCCCTGCGAGTGCCTGCATCCCTGTCATGGGGCCCTG 780

Qy  781 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTTT 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGGTGTATTATTTT 840

Qy  841 TCAAGCCCCTCCTTCCTTCCCAAATTCACAACAGCTCAAAATCTGCAGTCGAAACCCAAG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TCAAGCCCCTCCTTCCTTCCCAAATTCACAACAGCTCAAAATCTGCAGTCGAAACCCAAG 900

Qy  901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960

Qy  961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTTCGATGGGCAATGGGATCCC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTTCGATGGGCAATGGGATCCC 1020

Qy 1021 CACATTGTTGAGTGGCAC 1038
        ||||||||||||||||||
Db 1021 CACATTGTTGAGTGGCAC 1038
```

```
RESULT 2
AAL53846
ID   AAL53846 standard; DNA; 1038 BP.
XX
AC   AAL53846;
XX
DT   13-FEB-2003  (first entry)
XX
DE   DNA of human G-Protein Coupled Receptor gene.
XX
KW   Anorectic; antidiabetic; antilipaemic; antiasthmatic; antiinflammatory;
KW   antiallergic; antianginal; nephrotropic; hepatotropic; immunosuppressive;
KW   virucide; G-Protein-agonist; G-Protein-antagonist; dyslipidaemia; GPCR;
KW   G-Protein Coupled Receptor; obesity; diabetes; asthma; bronchitis;
KW   allergy; angina; glomerulonephritis; hepatitis; allograft rejection;
KW   human; gene; ds.
XX
OS   Homo sapiens.
XX
FH                    Location/Qualifiers
FT   CDS              1..1038
FT                    /*tag= a
FT                    /partial
FT                    /product= "Human GPCR protein"
FT                    /note= "No stop codon"
XX
PN   WO200283736-A2.
XX
PD   24-OCT-2002.
XX
PF   14-FEB-2002; 2002WO-US004397.
XX
PR   14-FEB-2001; 2001US-0269040P.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Elliott SG, Rogers N, Busse LA;
XX
DR   WPI; 2003-075524/07.
DR   P-PSDB; AAO26511.
XX
PT   New GPCR polypeptide and encoding nucleic acid molecule, useful for
PT   diagnosis, treatment and/or prevention of diseases associated with GPCR
PT   polypeptides, such as obesity, diabetes, asthma, allergies, angina and
PT   hepatitis.
XX
PS   Claim 1; Fig 1; 122pp; English.
XX
CC   The invention relates to a novel isolated G-Protein Coupled Receptor
CC   (GPCR) protein. The methods and compositions of the present invention are
CC   useful for diagnosis, treatment, amelioration and/or prevention of
CC   diseases associated with G-Protein Coupled Receptor (GPCR) polypeptides,
CC   such as obesity, diabetes, dyslipidaemia, asthma, bronchitis, allergies,
CC   angina, glomerulonephritis, hepatitis and allograft rejection. This
CC   polynucleotide sequence represents the DNA encoding the human GPCR
CC   protein of the invention
XX
SQ   Sequence 1038 BP; 207 A; 294 C; 277 G; 260 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1038; DB 7;  Length 1038;
Best Local Similarity   100.0%;  Pred. No. 5e-287;
Matches 1038;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
RESULT 3
AAH51008
ID   AAH51008 standard; DNA; 1041 BP.
XX
AC   AAH51008;
XX
DT   28-AUG-2001 (first entry)
XX
DE   Human nGPCR11 coding sequence #2.
XX
KW   G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW   signal transduction; schizophrenia; thyroid disorder; renal failure;
KW   rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW   cardiovascular disease; proliferative disorder; hormonal disorder;
KW   neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW   attention deficit-hyperactivity disorder/attention deficit disorder;
KW   Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW   rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW   neuroprotective; ds.
XX
OS   Homo sapiens.
XX
PN   WO200136473-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US031581.
XX
PR   16-NOV-1999;   99US-0165838P.
PR   17-NOV-1999;   99US-0166071P.
PR   19-NOV-1999;   99US-0166678P.
PR   28-DEC-1999;   99US-0173396P.
PR   22-FEB-2000;   2000US-0184129P.
PR   28-FEB-2000;   2000US-0185421P.
PR   28-FEB-2000;   2000US-0185554P.
PR   02-MAR-2000;   2000US-0186530P.
PR   03-MAR-2000;   2000US-0186811P.
PR   09-MAR-2000;   2000US-0188114P.
PR   17-MAR-2000;   2000US-0190310P.
PR   21-MAR-2000;   2000US-0190800P.
PR   20-APR-2000;   2000US-0198568P.
PR   02-MAY-2000;   2000US-0201190P.
PR   08-MAY-2000;   2000US-0203111P.
PR   25-MAY-2000;   2000US-0207094P.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Vogeli G, Wood LS,  Parodi LA,  Hiebsch RR,  Lind P,  Slightom J;
PI   Schellin KA,  Kaytes PS,   Bannigan CM,   Ruff V,   Sejlitz T,  Huff RM;
XX
DR   WPI; 2001-389825/41.
DR   P-PSDB; AAG80968.
XX
PT   New G protein-coupled receptor (nGPCRx) and its encoding polynucleotide
PT   useful for diagnosing and treating e.g. schizophrenia.
XX
PS   Claim 4; Page 89; 261pp; English.
XX
CC   The present invention relates to novel G protein-coupled receptors
CC   (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
CC   31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
CC   the coding sequence for one such G protein-coupled receptor. GPCRs are
CC   also known as seven transmembrane receptors and function in signal
CC   transduction. The nGPCRx coding sequences are useful for screening a
CC   human to diagnose a disorder affecting the brain or a genetic
CC   predisposition, specifically schizophrenia. nGPCRx are useful for
CC   identifying compounds useful for treating schizophrenia. Detection of
CC   nGPCRx in a sample is useful as a diagnostic tool for diseases or
CC   disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
CC   CNS disorders, infections such as HIV-1, metabolic and cardiovascular
CC   diseases, proliferative disorders and hormonal disorders. Modulators of
CC   nGPCRx activity have the utility for treating neurological disorders,
```

```
Qy      1 ATGTACAACGGGTCGTGCTGCCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG   60
          ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTACAACGGGTCGTGCTGCCGCCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG   60

Qy     61 CGCGTCATGTGGCCTTTGTGCTGGCGGCACTAGGCGCAATGGGCGTCGCCCGTGTGGTTTC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCGTCATGTGGCCTTTGTGCTGGCGGCACTAGGCGCAATGGGCGTCGCCCGTGTGGTTTC  120

Qy    121 TGCTTCCACATGAAGCCTGGAAGCCACTGTTTACCTTTCAATTGGCCGTGGCT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTTCCACATGAAGCCTGGAAGCCACTGTTTACCTTTCAATTGGCCGTGGCT  180

Qy    181 GATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240

Qy    241 TGGGCTTTGGGACATTCCTGCGAGTGGGGCTCTCACGTTGGCCATGAACAGGGCC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTTGGGACATTCCTGCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300

Qy    301 GGGAGCATGTCGTGTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATGTCGTGTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC  360

Qy    361 CACCACGGCGTGAACACTATCTCCACCCGGTGGCGCCTGGCATCGTCTGCACCCTGTG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGGCGTGAACACTATCTCCACCCGGTGGCGCCTGGCATCGTCTGCACCCTGTGG  420

Qy    421 GCCTGGTCATCCTCGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCTGGTCATCCTCGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480

Qy    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGGCTGGCATGACATCATG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGGCTGGCATGACATCATG  540

Qy    541 TTTCCAGTTCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTTCCAGTCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600

Qy    601 TGGAGCCTGAGGCCGGAGCAGCTGGCCAGAGGCTCGGATGAAGAGAGGCGACCCGG  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGGCCGGAGCAGCTGGCCAGAGGCTCGGATGAAGAGAGGCGACCCGG  660

Qy    661 TTCATCATGGTGGTGGCAATGTGTTCATCACATGTGTTCATCACATGCCCAGCCTGTCTGCTGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGGTGGCAATGTGTTCATCACATGTGTTCATCACATGCCCAGCCTGTCTGCTGA  720

Qy    721 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCTTCGATCCCTCTGTCCATGGGGCCTG  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCTTCGATCCCTCTGTCCATGGGGCCTG  780

Qy    781 CACATAACCCTCAGCTTCACCTACACATGAACCACAGCATGCTCGATCCCCTGGTGTATTTT  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCAGCTTCACCTACACATGAACCACAGCATGCTCGATCCCCTGGTGTATTTT  840

Qy    841 TCAAGCCCCTCCTTTCCACAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGCCCCTCCTTTCCACAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG  900

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC  960

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1020

Qy   1021 CACATTGTTGAGTGGCAC  103B
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC  103B
```

```
CC    including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
CC    disorder/attention deficit disorder), and neuronal disorders such as
CC    Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
CC    Additional disorders include inflammatory conditions (e.g. Crohn's
CC    disease), rheumatoid arthritis, autoimmune disorders, cancers,
CC    respiratory ailments such as asthma, and inflammatory diseases e.g.
XX    inflammatory bowel disease
SQ    Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match    100.0%;  Score 1038;  DB 4;  Length 1041;
Best Local Similarity  100.0%;  Pred. No. 5e-287;
Matches 1038;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGTACAACGGGTCGTGCTGCCGCATCCCAGGTGATGCCGCCG   60
          |||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTACAACGGGTCGTGCTGCCGCATCCCAGGTGATGCCGCCG   60

Qy     61 CTGCTCATTGTGGCCTTGTGCTGGGCCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCTCATTGTGGCCTTGTGCTGGGCCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC  120

Qy    121 TGCTTCCACATGAAGACCTGGAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTTCCACATGAAGACCTGGAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  180

Qy    181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGACAGACTATTACCTCAGACGTAGACAC  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGACAGACTATTACCTCAGACGTAGACAC  240

Qy    241 TGGGCTTTTGGGGACAATTCCCGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTTTGGGGACAATTCCCGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCCGGACAGTATTCAAAGTGGTCCACCCC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCCGGACAGTATTCAAAGTGGTCCACCCC  360

Qy    361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCCATCGTCTCGCACCCTGTGG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCCATCGTCTCGCACCCTGTGG  420

Qy    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG  480

Qy    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGCCAATGACATCATG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGCCAATGACATCATG  540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTCCGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTCCGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600

Qy    601 TGGAGCCTGAGGCGGGAGCAGCTGGCCAGACAGGCCGGATGAAGAGGCGACCCGG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGGCGGGAGCAGCTGGCCAGACAGGCCGGATGAAGAGGCGACCCGG  660

Qy    661 TTCATCATGGTGGTGGCAATTGTGTTCATCACGAGCTACACCTGCCTGCCTACCTGTCTGCTAGA  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGGTGGCAATTGTGTTCATCACGAGCTACACCTGCCTGCCTACCTGTCTGCTAGA  720

Qy    721 CTCTATTTCCTCTGACGGTGCCCTGCCTCGAGTGCCCTGCCTCCTGTCTCATGGGGCCCTG  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGACGGTGCCCTGCCTCGAGTGCCCTGCCTCCTGTCTCATGGGGCCCTG  780

Qy    781 CACATAACCCTCAGCTTCACCTACGTTGCTGAACAGCAATGCTGGATCCCCTGGTGTATTATTT  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCAGCTTCACCTACGTTGCTGAACAGCAATGCTGGATCCCCTGGTGTATTATTT  840

Qy    841 TCAAGCCCCTCCTTTCCAAAATCTACAACAAGCTCACAGACAAGTCTGAAACCCAAG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGCCCCTCCTTTCCAAAATCTACAACAAGCTCACAGACAAGTCTGAAACCCAAG  900
```

```
Qy    901 CAGGCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CAGGCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC   960

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGGCCAGTCGATGGGCAATGGGATCCC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGGCCAGTCGATGGGCAATGGGATCCC  1020

Qy   1021 CACATTGTTGAGTGGCAC  1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC  1038
```

```
RESULT 4
AAS12581
ID   AAS12581 standard; cDNA; 1041 BP.
XX
AC   AAS12581;
XX
DT   19-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human G protein-coupled receptor (GPCR).
XX
KW   Human; G-protein coupled receptor; GPCR; chemokine receptor; protease;
KW   hyperproliferative disorder; neurological disorder; psychiatric disease;
KW   inflammatory disorder; respiratory disorder; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1041
FT                   /*tag= a
FT                   /product= "GPCR protein"
XX
PN   WO20017029-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US009522.
XX
PR   27-MAR-2000; 2000US-0192419P.
PR   06-SEP-2000; 2000US-0230459P.
PR   20-SEP-2000; 2000US-00666535.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Ye J, Cravchik A,  Di Francesco V,  Beasley EM;
XX
DR   WPI; 2001-616503/71.
DR   P-PSDB; AAU06197.
XX
PT   Novel human G-protein coupled receptor proteins and nucleic acid
PT   molecules encoding the protein for use in developing human therapeutics
PT   and diagnostic compositions and for identifying modulators of the
PT   protein.
XX
PS   Claim 23; Fig 1; 66pp; English.
XX
CC   The present invention relates to the isolation of a novel human G-protein
CC   coupled receptor (GPCR) which is related to the chemokine receptor
CC   subfamily. The cDNA and gene sequences encoding for GPCR are also given
CC   in the invention. The sequences of the invention are useful for
CC   diagnosing and treating diseases or conditions mediated by human
CC   proteases. Such diseases include hyperproliferative disorders (e.g.
CC   hyperplasia), neurological disorders (e.g. Parkinson's disease),
CC   psychiatric diseases (e.g. schizophrenia), inflammatory disorders (e.g.
CC   diabetes) and respiratory disorders (e.g. adult respiratory distress
CC   syndrome; ARDS). The GPCR protein is also useful for identifying a
CC   modulator of the expression of the protein. It also serves as a target
CC   for identifying agents for use in mammalian therapeutic applications,
CC   e.g. a human drug, particularly modulating a biological or pathological
CC   response in a cell or tissue that expresses the protein, in biological
```

CC   assays related to GPCRs that are related to members of the chemokine
CC   receptor subfamily, in drug screening assays and in competition binding
CC   assays. GPCR is also useful in diagnosing a disease or predisposition to
CC   a disease mediated by the peptide, in pharmacogenomic analysis. The
CC   polynucleotide sequences can also be used in gene therapy. The present
CC   sequence encodes for the novel human GPCR of the invention
XX
SQ   Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match   100.0%; Score 1038; DB 5; Length 1041;
Best Local Similarity 100.0%; Pred. No. 5e-287;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTCGTGCTGCCCATCGAGGGGACACCATCTCCCAGTGATGCCGCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCCATCGAGGGGACACCATCTCCCAGTGATGCCGCCG 60

Qy   61 CTGCTCATTGTGCCCTTTGTGCTGGGCCGCACTAGGCAATGGGGTGCGCCCTGTGTGGTTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGCCCTTTGTGCTGGGCCGCACTAGGCAATGGGGTGCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGAGTAGACAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGAGTAGACAC 240

Qy   241 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTTCACGTGGCCATGAACAGGGCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTTCACGTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGTGCTCGGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTTCCTTACGGTGGTGCTCGGCGGACAGGTATTTCAAAGTGGTCCACCCC 360

Qy   361 CACCACGCGGGTGAACACTATCTCCACCCGGGTGGCGGCTGCATCGTCTGCACCCGTGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGGGTGAACACTATCTCCACCCGGGTGGCGGCTGCATCGTCTGCACCCGTGG 420

Qy   421 GCCCTGGTCATCCTGGGAAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGTGCAAGAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGGAAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGTGCAAGAG 480

Qy   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGAGTCAGACATCATG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGAGTCAGACATCATG 540

Qy   541 TTCCAGCTGGAGTTCTTATGCCCCTGCATCATCTATTTGCTCCTTCAAGATTGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTATGCCCCTGCATCATCTATTTGCTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCGGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCGGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG 660

Qy   661 TTCATCATGGTGGTGCAATTGTGTTCATCACATGCACCTACCGTGTCTGCTGCTAGA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCATCATGGTGGTGCAATTGTGTTCATCACATGCACCTACCGTGTCTGCTGCTAGA 720

Qy   721 CTCATTTCCTCTGACGGTGCCCTCGAGTCCGAGTCCATCCCTCGTCCATGGGGCCCTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTCATTTCCTCTGACGGTGCCCTCGAGTCCGAGTCCATCCCTCGTCCATGGGGCCCTG 780

Qy   781 CACTAAACCCTCAGCTTCACCTACTGAACAGCATGCTGGATCCCCGGTGTATTATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CACTAAACCCTCAGCTTCACCTACTGAACAGCATGCTGGATCCCCGGTGTATTATTTT 840

Qy   841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTTCAAATCTGAGTCTGAAACCCAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTTCAAATCTGAGTCTGAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTGAACCTCGGTCGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGGATCCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
              ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 5
AAS07946
ID   AAS07946 standard; cDNA; 1041 BP.
XX
AC   AAS07946;
XX
DT   23-OCT-2001 (first entry)
XX
DE   Human cDNA encoding G-protein coupled receptor, hRUP19.
XX
KW   Human; G-protein coupled receptor; GPCR; hRUP19; agonist;
KW   inverse agonist; lung cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1041
FT   /*tag= a
FT   /product= "hRUP19"
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US031509.
XX
PR   17-NOV-1999; 99US-0166088P.
PR   17-NOV-1999; 99US-0166099P.
PR   17-NOV-1999; 99US-0166369P.
PR   23-DEC-1999; 99US-0171900P.
PR   23-DEC-1999; 99US-0171901P.
PR   23-DEC-1999; 99US-0171902P.
PR   11-FEB-2000; 2000US-0181749P.
PR   14-MAR-2000; 2000US-0189258P.
PR   14-MAR-2000; 2000US-0189259P.
PR   10-APR-2000; 2000US-0195898P.
PR   10-APR-2000; 2000US-0195899P.
PR   10-APR-2000; 2000US-0196078P.
PR   28-APR-2000; 2000US-0200419P.
PR   12-MAY-2000; 2000US-0203630P.
PR   12-JUN-2000; 2000US-0210741P.
PR   12-JUN-2000; 2000US-0210982P.
PR   21-AUG-2000; 2000US-0226760P.
PR   26-SEP-2000; 2000US-0235418P.
PR   26-SEP-2000; 2000US-0235779P.
PR   20-OCT-2000; 2000US-0242332P.
PR   20-OCT-2000; 2000US-0242343P.
PR   24-OCT-2000; 2000US-0243019P.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R, Dang HT, Lowitz KP;
XX
DR   WPI; 2001-355616/37.
DR   P-PSDB; AAI04373.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents.
XX

PS   Claim 47; Page 110; 160pp; English.
XX
CC   The sequence encodes a human G-protein coupled receptor (GPCR), hRUP19
CC   The endogenous and non-endogenous, constitutively activated versions of
CC   human G-protein coupled receptors (GPCR), are useful for direct
CC   identification of candidate compounds as receptor agonists, inverse
CC   agonists or partial agonists having applicability as therapeutic agents
CC   for treating diseases related to GPCR, e.g. lung cancer. Non-endogenous
CC   version of human GPCRs are also utilized in research settings and in
CC   vitro and in vivo system, incorporating GPCRs can be utilised to
CC   elucidate and understand the roles these receptors play in the human
CC   condition, both normal and diseased
XX
SQ   Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match   100.0%; Score 1038; DB 5; Length 1041;
Best Local Similarity   100.0%; Pred. No. 5e-287;
Matches 1038;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATGAGGGGGACACCATCTCCCAGTGATGCGCGCCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGTGATGCGCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTGTGCTGGTGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTGTGCTGGTGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCAGGAGCCAGCACTGTTTACCTTTCAATTTGGCGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCAGGAGCCAGCACTGTTTACCTTTCAATTTGGCCGGCT 180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCCTGCGAGTGGGGTCTTCACGTTGGCCATGAACAGGGGCC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCCTGCGAGTGGGGTCTTCACGTTGGCCATGAACAGGGGCC 300

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGGCGGACAGGTATTTCAAAGTGGTCCACCCC 360

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGCTGGCATCGTCTGCACCCTGTGG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGGTGAACACTATCTCCACCCGGGCTGGCATCGTCTGCACCCTGTGG 420

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCGGCAGCAGTCGGAGCAGAGGACGGTCGCCCGG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCGGCAGCAGTCGGAGCAGAGGACGGTCGCCCGG 660

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATCTACCTGCCCAGCGTGTTCTGCTAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATCTACCTGCCCAGCGTGTTCTGCTAGA 720

Qy   721 CCTATTTCCTCTGACGCGGTGCCCTGCGATCCCTGTCCATGGGCGCCCTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCTATTTCCTCTGACGCGGTGCCCTGCGATCCCTGTCCATGGGCGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGATCCCCTGGTGTGTATTATTT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   781 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGATCCCCTGGTGTGTATTATTT 840

Qy   841 TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCGCAGTCTGAGAACCCAAG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAACAAATCGCAGTCTGAAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
           ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 6
AAD27497
ID   AAD27497 standard; DNA; 1041 BP.
XX
AC   AAD27497;
XX
DT   18-APR-2002 (first entry)
XX
DE   Human G-protein coupled receptor (GPCRx14) DNA.
XX
KW   Human; G-protein coupled receptor; GPCRx14; cerebroprotective; vomiting;
KW   receptor-mediated disorder; therapy; urinary retention; allergy; obesity;
KW   osteoporosis; angina pectoris; restenosis; atherosclerosis; hypotension;
KW   anorexia; tumour; migraine; acute heart failure; ulcer; antiinflammatory;
KW   stroke; hypertension; neuronal disorder; myocardial infarction psychotic;
KW   depression; mental retardation; neurodegenerative disease; antibacterial;
KW   Alzheimer's disease; dementia; ischaemia; Parkinson's disease; antiviral;
KW   Huntington's disease; anxiety; antifungal; immunosuppressive; cytostatic;
KW   vulnerary; analgesic; anorectic; anabolic; diuretic; cardiant; nootropic;
KW   antiemetic; vasotropic; diabetes; cancer; tranquilliser; neuroleptic; ds
XX
OS   Homo sapiens.
XX
FH   Location/Qualifiers
FT   Key
FT   CDS   1..1041
FT   /*tag= a
FT   /product= "Human GPCRx14 protein"
XX
PN   WO200198330-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-BE000104.
XX
PR   20-JUN-2000; 2000US-0212913P.
PR   11-JUL-2000; 2000US-0217494P.
PR   26-JAN-2001; 2001EP-00870015.
PR   12-FEB-2001; 2001EP-00870024.
XX
PA   (EURO-) EUROSCREEN SA.
XX
PI   Lannoy V, Brezillon S, Detheux M, Parmentier M, Govarts C;
XX
DR   WPI; 2002-130789/17.
DR   P-PSDB; AAE17077.
XX
PT   New G-protein coupled receptor, useful in the manufacture of medicaments
PT   for treating receptor mediated disorders e.g. acute heart failure and
PT   Alzheimer's disease.
XX
PS   Example 2; Page 28-29; 46pp; English.
XX
CC   The present invention relates to a G-protein coupled receptor (GPCR) and
CC   nucleotide encoding it. GPCR are useful in the manufacture of a

CC medicament for the prevention and/or treatment of receptor-mediated
CC disorders e.g. viral infections, virus and bacterial diseases, diseases
CC and disorders involving disturbances of cell migration, diseases or
CC perturbations of immune system including cancers; development of tumours
CC and tumour metastasis, inflammatory and neoplastic processes; bacterial
CC and fungal infections, in wound and bone healing, dysfunction of
CC regulatory growth functions; pains, diabetes, obesity, anorexia, bulimia,
CC urinary retention, osteoporosis, angina pectoris, atherosclerosis,
CC restenosis, diseases involving excessive or reduced proliferation or loss
CC of smooth muscle cells, aneurysm, stroke, ischaemia, ulcers, allergies,
CC benign prostatic hypertrophy, migraine, vomiting; blood circulating
CC affections including acute heart failure, hypotension, hypertension and
CC myocardial infarction psychotic; neuronal disorders such as anxiety,
CC schizophrenia, maniac depression, depression, delirium, dementia, severe
CC mental retardation; degenerative diseases; neurodegenerative diseases
CC such as Alzheimer's disease, Parkinson's disease; and dyskinesias e.g.
CC Huntington's disease or Gilles de la Tourette's syndrome and other
CC related diseases. The present sequence is GPCRx14 DNA
XX
SQ Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match   100.0%; Score 1038; DB 6; Length 1041;
Best Local Similarity 100.0%; Pred. No. 5e-287;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTCGTGCGCGCATCCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCGCGCATCCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60

Qy   61 CTGCGTCCATTGTGGCCTTTGTGCTGGGCGCACTAGGACAGCCAATGGGGTCGCCCTGTGTTC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCGTCCATTGTGGCCTTTGTGCTGGGCGCACTAGGACAGCCAATGGGGTCGCCCTGTGTTC 120

Qy 121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTTACCTTTCAATTGGCCGTGGCT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTTACCTTTCAATTGGCCGTGGCT 180

Qy 181 GATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy 241 TGGGCTTTTGGGGACTCCCGCCGAGTGGGCGCTCCTCACGTTGGCCATGAACAGGGCC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 TGGGCTTTTGGGGACTCCCGCCGAGTGGGCGCTCCTCACGTTGGCCATGAACAGGGCC 300

Qy 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGTGTATTTCAAAGTTGGTCCACCCC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGTGTATTTCAAAGTTGGTCCACCCC 360

Qy 361 CACCACGCCGTGAACACTATCTCCACCCGGGTGGGCGGTCGCATCGTCTCGCACCCTGTGG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 CACCACGCCGTGAACACTATCTCCACCCGGGTGGGCGGTCGCATCGTCTCGCACCCTGTGG 420

Qy 421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy 481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540

Qy 541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTGCTCCTTCAAGATTGTT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTGCTCCTTCAAGATTGTT 600

Qy 601 TGGAGCCTGAGGCGGAGGCAGCAGCAGCGGCCAGACAGGCTCGGATGAAGAGGCGACCCG 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 TGGAGCCTGAGGCGGAGGCAGCAGCAGCGGCCAGACAGGCTCGGATGAAGAGGCGACCCG 660

Qy 661 TTCATCACGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTCGTCTGCTAGA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 TTCATCACGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTCGTCTGCTAGA 720

Qy 721 CTCTATTCCTCTGGACGGTGCCTGCGATCCTCGTCCATGGGGCCCTG 780
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 721 CTCTATTCCTCTGGACGGTGCCTGCGATCCTCGTCCATGGGGCCCTG 780

Qy 781 CACATAACCCTCAGCTTCACCTACAGACAGCACGATGCTGGATCCCCTGGTGTATTATTT 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 CACATAACCCTCAGCTTCACCTACAGACAGCACGATGCTGGATCCCCTGGTGTATTATTT 840

Qy 841 TCAAGCCCCTCCTTCCAAATTCTACAACAAGCTCAAAATCTGCAGTCAGTCGAAACCCAAG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 TCAAGCCCCTCCTTCCAAATTCTACAACAAGCTCAAAATCTGCAGTCAGTCGAAACCCAAG 900

Qy 901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960

Qy 961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy 1021 CACATTGTTGAGTGGCAC 1038
         ||||||||||||||||||
Db 1021 CACATTGTTGAGTGGCAC 1038

RESULT 7
ABS70241
ID ABS70241 standard; DNA; 1041 BP.
XX
AC ABS70241;
XX
DT 26-NOV-2002 (first entry)
XX
DE DNA encoding human G protein-coupled receptor, nGPCR-11.
XX
KW Human; gene; ds; G protein-coupled receptor; GPCR; nGPCR; beGPCR;
KW nG protein coupled receptor; communication; serpentine structure;
KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
KW genetic predisposition; brain; immune response; gene therapy;
KW anxiety disorder; depression; bipolar disorder; schizophrenia;
KW Huntington's disease; dyskinesia; manic depression; stroke;
KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN WO200264789-A1.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2001; 2001WO-US004641.
PR 14-FEB-2001; 2001WO-US004641.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lind P, Parodi LA, Vogeli G, Wood LS;
XX
DR WPI; 2002-674879/72.
DR P-PSDB; ABG93786.
XX
PT New nucleic acids and polypeptides of the nG protein-coupled receptor,
PT useful for treating or diagnosing a mental disorder or a disorder
PT affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
PT Parkinson's disease.
XX
PS Example 1; Page 84; 244pp; English.
XX
CC The invention discloses an isolated human polypeptide, and encoding
CC nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
CC nG protein coupled receptor-14 (nGPCR-14). GPCRs are vital in the
CC communication between cells and their environment and are characterised
CC by a serpentine structure that passes through the cell membrane seven

CC   times, hence the reason such receptors are sometimes called seven
CC   transmembrane receptors (7TM). The polynucleotides and polypeptides are
CC   useful for identifying an nGPCR allelic variant that correlates with a
CC   mental disorder, for isolating an antibody that binds to an epitope of
CC   the polypeptide, for identifying a compound that binds the polypeptide or
CC   polynucleotide and/or modulates its biological activity, for screening a
CC   human subject to diagnose a disorder, or a genetic predisposition to a
CC   disorder, affecting the brain or a genetic disposition to the disorder,
CC   for identifying compounds useful for the treatment of a mental disorder
CC   and for identifying a compound useful as a modulator of binding between
CC   nGPCR-14 and a binding partner of nGPCR-14. The polypeptide is also
CC   useful for inducing an immune response in a mammal. The nucleic acid or
CC   polypeptide is particularly useful, using gene therapy, for treating e.g.
CC   anxiety disorders, depression, bipolar disorder, schizophrenia,
CC   Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
CC   disease or Alzheimer's disease. The nucleic acid and polypeptide may also
CC   be used for treating diabetes, inflammation or wounds. The sequences
CC   presented in ABS70202-ABS70249, ABS70338 and ABS70243 are the DNAs
CC   encoding the nGPCRs (also referred to as beGPCRs)
XX
SQ   Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match   100.0%; Score 1038; DB 6; Length 1041;
Best Local Similarity   100.0%; Pred. No. 5e-287;
Matches 1038; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCGCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCGCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60

Qy   61 CTGCTCATTGTGTGCCTTTGTGCTGGGCGCCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGTGCCTTTGTGCTGGGCGCCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTACCTTTTCAATTGGCCGTGGCT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTACCTTTTCAATTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTATGATCGCGCCTTTCGGACAGAGTATTACCTCAGACGACTAGACAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTATGATCGCGCCTTTCGGACAGAGTATTACCTCAGACGACTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCGCTCTTCACGTTGGCCATGAACAGGGCC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCGCTCTTCACGTTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATCGTGTCCTTACGGTGGTGCTGCGACCAGGTATTTCAAAGTGGTCCACCCC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTCCTTACGGTGGTGCTGCGACCAGGTATTTCAAAGTGGTCCACCCC 360

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGGCGCATCGTCTGCTGCACCCCTGTGG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGGCGCATCGTCTGCTGCACCCCTGTGG 420

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGGTGCAAGAG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGGTGCAAGAG 480

Qy   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGCTCGGCCAATGGCTGGCATGACATCATG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGCTCGGCCAATGGCTGGCATGACATCATG 540

Qy   541 TCCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTATTTTTGCTCCTTCAAGATTGTT 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTATTTTTGCTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCGCAGCTCGCCAGACAGCTCGGATGAGAAGGCGACCCGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCGCAGCTCGCCAGACAGCTCGGATGAGAAGGCGACCCGG 660

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACAATGGTCACCTGCCAGCGTGTCTGCTAGA 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACAATGGTCACCTGCCCAGCGTGTCTGCTAGA 720

Qy   721 CTCTATTTCCTTCTGGACGGTGCCCTGCCCTGGAGTGCCTGCGATCCGATCCCGTGTCCATGGGCCCTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTCTATTTCCTTCTGGACGGTGCCCTGCCCTGGAGTGCCTGCGATCCGATCCCGTGTCCATGGGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACTACTACAGCAATGCTGATCCCTGGTGTAGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CACATAACCCTCAGCTTCACTACTACAGCAATGCTGATCCCTGGTGTAGATTTT 840

Qy   841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAAATCTGAGTCTGAAACCCAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAAATCTGAGTCTGAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCAGTGGATGGGCAATGGGCAATGGGATCCC 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGTTGCATCAGTCAGTGTGGCAAAATAGTTTCCAAAGCCAGTCAGTGGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
              ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 8
ABZ42873
ID   ABZ42873 standard; DNA; 1041 BP.
XX
AC   ABZ42873;
XX
DT   06-MAR-2003 (first entry)
XX
DE   Human GPCR polynucleotide SEQ ID NO 7.
XX
KW   Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
KW   drug development; gustatory; taste; fragrance; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200216548-A2.
XX
PD   28-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-IB001446.
XX
PR   04-AUG-2000; 2000JP-00237818.
PR   13-FEB-2001; 2001JP-00034434.
XX
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI   Haga T, Takeda S, Mitaku S;
XX
DR   WPI; 2002-304118/34.
DR   P-PSDB; ABP95599.
XX
PT   Database global search for G protein-coupled receptors, proteins and
PT   encoded genes for studying in vivo signal transduction mechanism and
PT   identifying targets for drug development.
XX
PS   Claim 9; SEQ ID NO 7; 97pp + Sequence Listing; Japanese.
XX
CC   The invention relates to a method for screening G protein-coupled
CC   receptor (GPCR) genes (ABZ48270-ABZ43216) and/or GPCR proteins (ABP95596-
CC   ABP95942) by extracting open-reading frames containing 6-8 transmembrane
CC   domains with 250-1000 amino acid residues to give a gene homologous with
CC   a known GPCR gene. The receptor proteins and encoded genes are useful for
CC   studying in vivo signal transduction mechanism and identifying targets
CC   for drug development e.g. based on olfactory and gustatory receptors in
CC   form of agonists and antagonists by screening intrinsic and extrinsic
CC   ligands as bitter taste inhibitors, taste enhancers and fragrance
CC   improvers. Note: The sequence data for this patent did not form part of

```
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1038;  DB 6;  Length 1041;
Best Local Similarity   100.0%;  Pred. No. 5e-287;
Matches 1038;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCCGGCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCCGGCCG 60

Qy         61 CTGCTCATTGTGGCCTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCATTGTGGCCTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120

Qy        121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTCGGCT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTCGGCT 180

Qy        181 GATTTCCTCTTCTATGATCTGCCTGCCCTTTCCGACAGACTATTACCTCAGAGTAGACAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATTTCCTCTTCTATGATCTGCCTGCCCTTTCCGACAGACTATTACCTCAGAGTAGACAC 240

Qy        241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACAGGGCC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACAGGGCC 300

Qy        301 GGGAGCAATCGTGTTCCTTACGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCCCC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGAGCAATCGTGTTCCTTACGGTGGCTGGCGACAGGTATTTCAAAGTGGTCCACCCCC 360

Qy        361 CACCACGCGGTGAACACTATCTCCACCGGGTGGCCGGCTGGTGCTGCTGCACCCTGTGG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CACCACGCGGTGAACACTATCTCCACCGGGTGGCCGGCTGGTGCTGCTGCACCCTGTGG 420

Qy        421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy        481 ACGGCCGTCTCCCTGTGAGAGCTTCATCATGAGTCGGCTGGCATGGCTGGCATCATG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACGGCCGTCTCCCTGTGAGAGCTTCATCATGAGTCGGCTGGCATGGCTGGCATCATG 540

Qy        541 TTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTATTTGCTCTTCAAGATTGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTATTTGCTCTTCAAGATTGTT 600

Qy        601 TGGAGCCTGAGGCGGAGGCGAGCAGCTGGCCAGACAGCTCGGATGAAGAAGAGGCACCCGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGGAGCCTGAGGCGGAGGCGAGCAGCTGGCCAGACAGCTCGGATGAAGAAGAGGCACCCGG 660

Qy        661 TTCATCATGGTGGTGCAATTGTGTTCATCAGAGTCGCCCAGCCGTCTGCTCGCTCGTAGA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTCATCATGGTGGTGCAATTGTGTTCATCAGAGTCGCCCAGCCGTCTGCTCGCTCGTAGA 720

Qy        721 CTCTATTTCCCTCGACGGTGCCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGGGCCCTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTATTTCCCTCGACGGTGCCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGGGCCCTG 780

Qy        781 CACATAACCCTCAGCTTCACCTCAGCAGCATGCTGGATCCCCTGGTGTATTATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACATAACCCTCAGCTTCACCTCAGCAGCATGCTGGATCCCCTGGTGTATTATTTT 840

Qy        841 TCAAGCCCCTCCTTCCAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCAAGCCCCTCCTTCCAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900

Qy        901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAAGATGCCAATTTCGAACCTCGGTCGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAAGATGCCAATTTCGAACCTCGGTCGC 960

Qy        961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy       1021 CACATTGTTGAGTGGCAC 1038
              ||||||||||||||||||
Db       1021 CACATTGTTGAGTGGCAC 1038

RESULT 9
AAL42499
ID   AAL42499 standard; DNA; 1041 BP.
XX
AC   AAL42499;
XX
DT   28-JUN-2002   (first entry)
XX
DE   Human purinergic-like G-protein coupled receptor (AXOR87) DNA sequence.
XX
KW   Human; gene; ds; purinergic-like G-protein coupled receptor; AXOR87;
KW   immunity; autoimmunity; inflammation; immunodeficiency;
KW   bacterial infection; fungal infection; viral infection;
KW   protozoa infection; cancer; diabetes; obesity; anorexia; bulimia; asthma;
KW   psoriasis; rheumatoid arthritis; osteoarthritis; psychotic disorder;
KW   neurological disorder; vaccine; chromosome 12q24.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1041
FT                   /*tag= a
FT                   /product= "Human purinergic-like G-protein coupled
FT                   receptor AXOR87"
XX
PN   GB2365868-A.
XX
PD   27-FEB-2002.
XX
PF   25-MAY-2001; 2001GB-00012860.
XX
PR   30-MAY-2000; 2000US-00580675.
PR   02-NOV-2000; 2000GB-00026839.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Ignar DM, Elshourbagy N, Gattu M, Shabon U;
XX
DR   WPI; 2002-364852/40.
DR   P-PSDB; AAO14788.
XX
PT   New purinergic-like G-protein coupled receptor AXOR87 polypeptide and
PT   polynucleotide, useful for treating diseases related to autoimmunity,
PT   inflammation, immunodeficiency, or bacterial, fungal, viral and protozoal
PT   infections.
XX
PS   Claim 5; Page 36; 47pp; English.
XX
CC   The invention comprises the amino acid and coding sequence of the human
CC   purinergic-like G-protein coupled receptor AXOR87. The AXOR87 DNA and
CC   protein sequences of the invention may be used for treating diseases
CC   related to immunity, autoimmunity, inflammation, immunodeficiency, and
CC   infections (i.e. bacterial, fungal, viral, protozoan). The AXOR87 DNA and
CC   protein sequences are particularly useful for treating: cancers,
CC   diabetes, obesity, anorexia, bulimia, asthma, psoriasis, rheumatoid
CC   arthritis, osteoarthritis, as well as psychotic and neurological
CC   disorders. The AXOR87 DNA and protein sequences may also be used as
CC   vaccines. The present DNA sequence (located on chromosome 12q24) encodes
CC   the human AXOR87 protein
XX
SQ   Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;
```

Query Match   100.0%; Score 1038; DB 6; Length 1041;
Best Local Similarity 100.0%; Pred. No. 5e-287;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGGTGATGCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGGTGATGCGCCG 60

Qy     61 CTGCTCCATTGTGGCCTTGTGTGGCGCACTAGGCGAATGGGGTCGCCCTGTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCTCCATTGTGGCCTTGTGTGGCGCACTAGGCGAATGGGGTCGCCCTGTGTGGTTTC 120

Qy    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180

Qy    181 GATTTCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATTTCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy    241 TGGGCTTTGGGACAATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTTGGGACAATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGGCC 300

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGCGGACAGGTATTTCAAAGTGGTCCACCCC 360

Qy    361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGGCTCGGCATGTCTCGCACCCTGTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGGCTCGGCATGTCTCGCACCCTGTGG 420

Qy    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGTGGAGAACCATCTCGGTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGTGGAGAACCATCTCGGTGCGTGCAAGAG 480

Qy    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600

Qy    601 TGGAGCCTGAGCGCGGAGGCGCAGCTGGCCAGACAGAGGCTCGGATGAAGAGGCGACCCG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGCGCGGAGGCGCAGCTGGCCAGACAGAGGCTCGGATGAAGAGGCGACCCG 660

Qy    661 TTCATCATGGTGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCTGTGTCTGCTACGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCTGTGTCTGCTACGA 720

Qy    721 CTCTATTTCCTCTGGAGCGTGCCCTCGCGATCCCGATCCCTGTCCATGGGGGCCCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGGAGCGTGCCCTCGCGATCCCGATCCCTGTCCATGGGGGCCCTG 780

Qy    781 CACATAACCCTCAGCTTCACTACACGATGCTGGATCCCGTGTATTATTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCAGCTTCACTACACGATGCTGGATCCCGTGTATTATTT 840

Qy    841 TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900

Qy    901 CAGCCAGGACACTCAAAAACACAAGGCCGGAAGAGATCGCCAATTTCGAACCTCGGTCGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CAGCCAGGACACTCAAAAACACAAGGCCGGAAGAGATCGCCAATTTCGAACCTCGGTCGC 960

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038
```

Db

RESULT 10
ABZ22648
ID   ABZ22648 standard; cDNA, 1041 BP.
XX
AC   ABZ22648;
XX
DT   31-MAR-2003 (first entry)
XX
DE   Human GAVE3 encoding cDNA SEQ ID NO:1.
XX
KW   Human; GAVE3; G protein-coupled receptor; GPCR; antiinflammatory;
KW   antiasthmatic; antirheumatic; antiarthritic; inflammatory disorder;
KW   asthma; chronic obstructive pulmonary disease; rheumatoid arthritis;
KW   chromosome 12; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT
FT   CDS   1..1041
FT   /*tag= a
FT   /product= "GAVE3"
FT   /note= "G protein-coupled receptor (GPCR)"
XX
PN   WO2003000846-A2.
XX
PD   03-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019490.
XX
PR   22-JUN-2001; 2001US-00886041.
XX
PA   (AVET) AVENTIS PHARM INC.
XX
PI   Xia T, Ni D, Bishingdrelo H, Ardati A, Minnich A, Jupp R;
XX
DR   WPI; 2003-184040/18.
DR   P-PSDB; ABP56751.
XX
PT   New GAVE3 nucleic acid and polypeptide, for preventing or treating a
PT   disorder associated with aberrant GAVE3 expression or activity, e.g.
PT   inflammatory disease, asthma, chronic obstructive pulmonary disease or
PT   rheumatoid arthritis.
XX
PS   Claim 1; Fig 1; 90pp; English.
XX
CC   The present sequence encodes a human G protein-coupled receptor (GPCR),
CC   designated GAVE3. GAVE3 is located on human chromosome 12. GAVE3 has
CC   antiinflammatory, antiasthmatic, antirheumatic, antiarthritic
CC   activities. GAVE3 nucleic acids, polypeptides, agonists and antagonists
CC   can be used for preventing or treating a disorder associated with
CC   aberrant GAVE3 expression or activity, e.g. inflammatory disorders such
CC   as asthma, chronic obstructive pulmonary disease or rheumatoid arthritis.
CC   The nucleic acids and polypeptides can also be useful for identifying
CC   modulators, i.e. candidate or test compounds, that bind to GAVE3 proteins
CC   or have a stimulatory or inhibitory effect on GAVE3 expression or
CC   activity. The nucleotide sequences can also be used for chromosomal
CC   mapping, tissue typing or forensic biology. Host cells comprising GAVE3
CC   can be used for producing non-human transgenic animals for studying the
CC   function and/or activity of GAVE3, or for identifying and/or evaluating
CC   modulators of GAVE3 activity
XX
SQ   Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Db

Qy

Eb

Eb

Qy

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Cy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

ID AB257957 standard; cDNA; 1041 BP.

XX

AC AB257957;

XX

DT 14-APR-2003 (first entry)

XX

DE Human respiratory chemokine receptor coding sequence.

XX

KW G-protein coupled receptor; GPCR; receptor; chemokine; human;
KW antiasthmatic; antiinflammatory; antitussive; vaccine; gene; ss.

XX

OS Homo sapiens.

XX

FH Key   Location/Qualifiers

FT CDS   1..1041
FT   /*tag= a
FT   /product= "Human G-protein coupled receptor"

XX

PN WO2003002604-A2.

XX

PD 09-JAN-2003.

XX

PF 25-JUN-2002; 2002WO-EP007021.

XX

PR 26-JUN-2001; 2001US-0300944P.

XX

PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX

PI Bhatia U, Jones CE, Bouhelal R, Seuwen K, Tenaillon L;

XX

DR WPI; 2003-210243/20.
DR N-PSDB; ABP58453.

XX

PT New polypeptide, useful for diagnosing or treating e.g., asthma, chronic
PT obstructive pulmonary disease, emphysema, chronic bronchitis or acute
PT respiratory distress syndrome.

XX

PS Claim 3; Page 42-43; 44pp; English.

XX

CC The present sequence is the coding sequence for a novel human G-protein
CC coupled receptor (GPCR) characterised as a respiratory chemokine
CC receptor. The receptor is expressed in respiratory tissues and tissues
CC related to monocyte/macrophage migration/activation, airway remodeling,
CC airway fibrosis, regulation of epithelial differentiation, regulation of
CC mucus hypersecretion, regulation of mucocilliary clearance, regulation of
CC inflammation, modulation of neutrophil, T-cell and eosinophil migration
CC and/or activation, and regulation of epithelial cell or mast cell
CC activation. GPCR polypeptides and polynucleotides of the invention may be
CC useful in treatment of asthma, chronic obstructive pulmonary disease,
CC emphysema, chronic bronchitis, acute respiratory distress syndrome, cough
CC and acute bronchitis. Methods of identifying agonist and antagonist
CC compounds that stimulate or inhibit the function or level of the novel
CC receptor are also provided. GPCR polynucleotides can also be used for
CC recombinant production of GPCR polypeptides, as diagnostic reagents, for
CC chromosome localisation studies, and as vaccines

XX

SQ Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match   100.0%; Score 1038; DB 7; Length 1041;
Best Local Similarity   100.0%; Pred. No. 5e-287;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATGGAGCCGGGGGACACCATCTCCCAGTGATGCCGCCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCGCATGGAGCCGGGGGACACCATCTCCCAGTGATGCCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGGCCACTAGGCAATGGGGTCGCCCTGTCGTGGTTTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGGCCACTAGGCAATGGGGTCGCCCTGTCGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTGCCTTTTCGGACAGACTATTACCTCAGAGTAGACAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy   241 TGGGCTTTGGGGACATTCCTGCCGAGTGGGCTCTTCACGTGGCCATGAACAGGGCC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGCTCTTCACGTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGGCGACAGTATTTCAAAGTGGTCCACCCC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTCGGCGACAGTATTTCAAAGTGGTCCACCCC 360

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCCTGGCATCGTCTGCACCCTGTGG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCCTGGCATCGTCTGCACCCTGTGG 420

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCCTGCCAGAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCCTGCCAGAG 480

Qy   481 ACGGGCCGTCTCCGTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCATGACATCATG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGGCCGTCTCCGTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCATGACATCATG 540

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTTGCTCCTTCAAGATTGTT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTTGCTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCAGCCAGCAGCTGGCAGGAGCAGGCTCGGATGAAGAAGGCGACCCG 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCAGCCAGCAGCTGGCAGGAGCAGGCTCGGATGAAGAAGGCGACCCGG 660

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy   721 CTCTATTCCTCTGGACGGTGCCCTCGAGTGCCCTGCGATCCCTCTGTCCATGGGGCCCTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTCTATTCCTCTGGACGGTGCCCTCGAGTGCCCTGCGATCCCTCTGTCCATGGGGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACCTACTACATGAACAGCATGCTGGATCCTGGTGTATTATTT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CACATAACCCTCAGCTTCACCTACTACATGAACAGCATGCTGGATCCTGGTGTATTATTT 840

Qy   841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCGAATTTCGAACCTCGGTCGC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy   961 AGGAGGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAAATAGTTGGCAAATGGGATCCC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAAATAGTTGGCAAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
           ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 11
AB257957

KW   Human; G-protein coupled receptor; GPCR; infection; neoplastic process;
KW   inflammation; myocardial infarction; atherosclerosis; angina pectoris;
KW   hypertension; osteoporosis; antibacterial; cytostatic; fungicide; pain;
KW   diabetes; cancer; virucide; analgesic; cardiant; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1041
FT   /*tag= a
FT   /product= "Human GPCR protein"
XX
PN   US2003108986-A1.
XX
PD   12-JUN-2003.
XX
PF   20-FEB-2002; 2002US-00079384.
XX
PR   21-JUN-2001; 2001US-00885453.
XX
PA   (EURO-) EUROSCREEN SA.
XX
PI   Communi D, Lannoy V, Brezillon S, Detheux M, Parmentier M;
PI   Govaerts C;
XX
DR   WPI; 2003-810852/76.
DR   P-PSDB; ABW00810.
XX
PT   Novel G-protein coupled receptor useful for treating viral infections,
PT   bacterial infections, fungal infections, cancer, diabetes, hypertension,
PT   osteoporosis, angina pectoris, myocardial infarction, atherosclerosis.
XX
PS   Claim 5; Fig 9; 0pp; English.
XX
CC   The present invention relates to novel G-protein coupled receptors
CC   (GPCRs) and the nucleic acids encoding them. The invention is useful for
CC   treating viral, bacterial and fungal infections, inflammatory and
CC   neoplastic processes, pain, diabetes, hypertension, osteoporosis, cancer,
CC   angina pectoris, myocardial infarction and atherosclerosis. The present
CC   sequence is human G-protein coupled receptor (GPCR) DNA
XX
SQ   Sequence 1041 BP; 208 A; 294 C; 278 G; 261 T; 0 U; 0 Other;

Query Match   100.0%; Score 1038; DB 9; Length 1041;
Best Local Similarity 100.0%; Pred. No. 5e-287;
Matches 1038;   Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTCGTGTCGCCGCATCGAGGGGACCACCTCCCAGTGATGCGGCCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGTCGCCGCATCGAGGGGACCACCTCCCAGTGATGCGGCCGCCG 60

Qy   61 CTGCTCAATTGTGCCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCGTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCAATTGTGCCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGAACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGAACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCTCGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTCGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATCGTGTTCCTTACCGTGGTGGCTCGCGGACAGGTATTTCAAAGTGGTCCACCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTTCCTTACCGTGGTGGCTCGCGGACAGGTATTTCAAAGTGGTCCACCC 360

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCCTGGCATCGTCTGCACCCTGTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCCTGGCATCGTCTGCACCCTGTGG 420

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCATGACATCATG 540

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTATTTTGCTCCTTCAAGATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTATTTTGCTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCAGCAGGCTGGCCAGACAGCTCGGATGAAGAAGGCGACCCGG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCAGCAGGCTGGCCAGACAGCTCGGATGAAGAAGGCGACCCGG 660

Qy   661 TTCATCATGGTGGTGCAATTGTGTTCATCACACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCATCATGGTGGTGCAATTGTGTTCATCACACATGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy   721 CTCTATTCCTCTGACGGTGCCCTCGGATGGCCTTGCCTGCCTCGTGTCCATGGGGCCCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTCTATTCCTCTGACGGTGCCCTCGGATGGCCTTGCCTGCCTCGTGTCCATGGGGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACCTACAATGAACAGCATGCTGGATCCCCTGGTGTATTATTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CACATAACCCTCAGCTTCACCTACAATGAACAGCATGCTGGATCCCCTGGTGTATTATTT 840

Qy   841 TCAAGCCCCCTCCTTTCCCAAATTCCACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCAAGCCCCCTCCTTTCCCAAATTCCACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAACACAAAGGCCGAAGAGATGCCAATTCGAACCTCGGTCGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGCCAGGACACTCAAAACACAAAGGCCGAAGAGATGCCAATTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 12
AAD61654
ID   AAD61654 standard; DNA; 1041 BP.
XX
AC   AAD61654;
XX
DT   15-JAN-2004 (first entry)
XX
DE   Human GPCRx14 DNA.
XX

```
Db   361  CACCACGCCGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG  420

Qy   421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGCGAGAAACCATCCTGCGTGCCAAGAG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGCGAGAAACCATCCTGCGTGCCAAGAG  480

Qy   481  ACGGCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCGATGACATCATG   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ACGGCGTCTCCCGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCGATGACATCATG   540

Qy   541  TTCCAGCTGAGTCTTTATGCCCCTGGCATCATCTTATTTTGCCCCTTCAAGATTGTT   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TTCCAGCTGAGTCTTTTATGCCCCCGGCATCATCATTTATTTTGCCCCTTCAAGATTGTT   600

Qy   601  TGGAGCCTGAGGCGGAGGCAGCAGCAGCAGCGGCTCGGATGGAGGAGGGCGACCCGG   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TGGAGCCTGAGGCGGAGGCAGCAGCAGCAGCGGCTCGGATGGAGGAGGGCGACCCGG   660

Qy   661  TTCATCATGTGGTGGCAATTGTGTTCATCACATGTACCCTGCCCAGCGTGTCTGCTAGA   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TTCATCATGTGGTGGTGCAATTGTGTTCATCACATGTACCCTGCCCAGCGTGTCTGCTAGA   720

Qy   721  CTCTATTCCTCTGACGGTGCCCTCGAGTGCCTCGCGATCCCTCGTCCATGGGGGCCCTG   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CTCTATTCCTCTGACGCGGTGCCCTCGAGTGCCTCGCGATCCCTCGTCCATGGGGGCCCTG   780

Qy   781  CACATAACCCTCAGCTTCACCTACATGACGAAGCACATGCTCGGATCCCTGGTGTATTATTT   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CACATAACCCTCAGCTTCACCTACATGACAGCACATGCTCGGATCCCTGGTGTATTATTTT   840

Qy   841  TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGTCCAAAATCTGCAGTCTGAAACCCAAG   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TCAAGCCCCTCCCTTCCCAAATTCTACAACAAGTCCAAAATCTGCAGTCTGAAACCCAAG   900

Qy   901  CAGCCAGGACACTCAAAAACACCAAAGGCCGGAAGATGCCAATTTCGACCTCGGTCGC   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CAGCCAGGACACTCAAAAACACCAAAGGCCGGAAGATGCCAATTTCGACCTCGGTCGC   960

Qy   961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1020

Qy  1021  CACATTGTTGAGTGGCAC  1038
          ||||||||||||||||||
Db  1021  CACATTGTTGAGTGGCAC  1038
```

RESULT 13
ABA81530
ID   ABA81530 standard; DNA; 1050 BP.
XX
AC   ABA81530;
XX
DT   28-JAN-2002 (first entry)
XX
DE   Human GPCRib polynucleotide SEQ ID NO 3.
XX
KW   Human; GCPR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
KW   anabolic; cytostatic; antiviral; gene therapy; cardiomyopathy; obesity;
KW   anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
KW   asthma; Alzheimer's disease; Parkinson's disorder; Huntington's disease;
KW   infection; human immunodeficiency virus; HIV; ds.
XX
OS   Homo sapiens.
XX
PN   WO200174904-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US010241.
XX
PR   31-MAR-2000; 2000US-0193664P.
PR   05-APR-2000; 2000US-0194614P.

PR   06-APR-2000; 2000US-0195063P.
PR   06-APR-2000; 2000US-0195066P.
PR   06-APR-2000; 2000US-0195067P.
PR   06-APR-2000; 2000US-0195068P.
PR   06-APR-2000; 2000US-0195069P.
PR   06-APR-2000; 2000US-0195070P.
PR   06-APR-2000; 2000US-0195510P.
PR   21-JUL-2000; 2000US-0219855P.
PR   27-JUL-2000; 2000US-0221284P.
PR   28-JUL-2000; 2000US-0221325P.
PR   11-AUG-2000; 2000US-0224588P.
PR   11-OCT-2000; 2000US-0239611P.
PR   18-JAN-2001; 2001US-0262508P.
PR   23-JAN-2001; 2001US-0263433P.
PR   23-JAN-2001; 2001US-0263604P.
PR   30-JAN-2001; 2001US-0265161P.
PR   29-MAR-2001; 2001US-0823172.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Majumder K,   Verret CAM,   Casman SJ,   Wolenc AR,   Spaderna SK;
PI   Padigaru M,   Mistinu VS,   Tchernev VT, Spytek KA,   Li L, Baumgartner JC;
PI   Gusev VY;
XX
DR   WPI; 2001-639351/73.
DR   P-PSDB; ABB44522.
XX
PT   New human G-protein coupled receptor X, GPCRX, polypeptide useful in
PT   treatment or prevention of GPCRX associated disorders e.g. cardiomyopathy
PT   or atherosclerosis, and to screen for antagonists and agonists useful
PT   therapeutically.
XX
PS   Claim 9; Page 10; 157pp; English.
XX
CC   The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC   encode G-coupled protein-receptor related polypeptides (ABB44521-
CC   ABB44543). The isolated polypeptide having a sequence differing by no
CC   more than 15 % of amino acid residues from one of 22 amino acid sequences
CC   (or mature forms of the sequences), fully defined in the specification
CC   and corresponding to human G-protein coupled receptor X (GPCRX)
CC   polypeptides. The polypeptides have potential cardiant,
CC   antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The
CC   polypeptides can be administered therapeutically, especially using gene
CC   therapy and expressing the encoding DNA in vivo, to treat or prevent
CC   GPCRX-associated disorders, especially in humans. For example, they can
CC   be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC   related to signal processing and metabolic pathway modulation (e.g.
CC   obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
CC   sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC   disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC   haematopoietic disorders, developmental diseases, neurological disorders,
CC   bacterial, fungal, protozoal and viral infections (e.g. with human
CC   immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically
CC   to determine the presence of or predisposition to a disease associated
CC   with altered levels of the polypeptide in mammals (especially humans) by
CC   detecting alterations in polypeptide expression levels relative to
CC   control samples. They are useful to identify agents binding polypeptide
CC   (e.g. cellular receptors or downstream effectors) and/or agents
CC   modulating cellular polypeptide expression or activity, useful as
CC   antagonists and agonists in disease treatment
XX
SQ   Sequence 1050 BP; 211 A; 298 C; 279 G; 262 T; 0 U; 0 Other;

```
Query Match             100.0%;  Score 1038;  DB 4;  Length 1050;
Best Local Similarity   100.0%;  Pred. No. 5e-287;
Matches  1038;          Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1  ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCAGTGATGCCGCCG  60
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6  ATGTACAACGGGTCGTGCTGCCGCATCCCATCGAGGGGGACACCATCTCCAGTGATGCCGCCG  65

Qy    61  CTGCTCATTGTGGCCCTTTGTGCTGCGGGCGCCACTAGGCCAATGGGGTCGCCGTGGTGGTTTC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

DT 28-JAN-2002 (first entry)
XX
DE Human GPCR1a polynucleotide SEQ ID NO 1.
XX
KW Human; GCPR; G-coupled protein-receptor; cardiant; antiarteriosclerotic;
KW anabolic; cytostatic; antiviral; gene therapy; cardiomyopathy; obesity;
KW anorexia; diabetes; osteoporosis; Crohn's disease; multiple sclerosis;
KW asthma; Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW infection; human immunodeficiency virus; HIV; ds.
XX
OS Homo sapiens.
XX
PN WO200174904-A2.
XX
PD 11-OCT-2001.
XX
PP 30-MAR-2001; 2001WO-US010241.
XX
PR 31-MAR-2000; 2000US-0193664P.
PR 05-APR-2000; 2000US-0194614P.
PR 06-APR-2000; 2000US-0195063P.
PR 06-APR-2000; 2000US-0195066P.
PR 06-APR-2000; 2000US-0195067P.
PR 06-APR-2000; 2000US-0195068P.
PR 06-APR-2000; 2000US-0195069P.
PR 06-APR-2000; 2000US-0195070P.
PR 06-APR-2000; 2000US-0195510P.
PR 21-JUL-2000; 2000US-0219855P.
PR 27-JUL-2000; 2000US-0221284P.
PR 28-JUL-2000; 2000US-0221325P.
PR 11-AUG-2000; 2000US-0224588P.
PR 11-OCT-2000; 2000US-0239613P.
PR 18-JAN-2001; 2001US-0262508P.
PR 23-JAN-2001; 2001US-0263433P.
PR 23-JAN-2001; 2001US-0263604P.
PR 30-JAN-2001; 2001US-0265161P.
PR 29-MAR-2001; 2001US-00823172.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Majumder K, Vernet CAM, Casman SJ, Wolenc AR, Spaderna SK,
PI Padigaru M, Mishnu VS, Tchernev VT, Spytek KA, Li L, Baumgartner JC;
PI Gusev VT;
XX
DR WPI; 2001-639351/73.
DR P-PSDB; ABB44522.
XX
PT New human G-protein coupled receptor X, GPCRX, polypeptide useful in
PT treatment or prevention of GPCRX associated disorders e.g. cardiomyopathy
PT or atherosclerosis, and to screen for antagonists and agonists useful
PT therapeutically.
XX
PS Claim 9; Page 7; 157pp; English.
XX
CC The invention relates to nucleic acid sequences (ABA81529-ABA81552) that
CC encode G-coupled protein-receptor related polypeptides (ABB44522-
CC ABB44543). The isolated polypeptide having a sequence differing by no
CC more than 15 % of amino acid residues from one of 22 amino acid sequences
CC (or mature forms of the sequences) fully defined in the specification
CC and corresponding to human G-protein coupled receptor X (GPCRX)
CC polypeptides. The polypeptides have potential cardiant,
CC antiarteriosclerotic, anabolic, cytostatic and antiviral activity. The
CC polypeptides can be administered therapeutically, especially using gene
CC therapy and expressing the encoding DNA in vivo, to treat or prevent
CC GPCRX-associated disorders, especially in humans. For example, they can
CC be used to treat/prevent cardiomyopathy, atherosclerosis, disorders
CC related to signal processing and metabolic pathway modulation (e.g.
CC obesity, anorexia), diabetes, osteoporosis, Crohn's disease, multiple
CC sclerosis, asthma, cancers, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disorder, Huntington's disease), immune disorders,
CC haematopoietic disorders, developmental diseases, neurological disorders,
CC bacterial, fungal, protozoal and viral infections (e.g. with human
CC immunodeficiency virus (HIV)-1 or HIV-2). They can be used diagnostically

Db 66 CTGCTCATTGTGCCTTTGTGCTGGGCGGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 125

Qy 121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTCGTTTACCTTTTCAATTTGGCGTGGCT 180
             ||||||||||||||||||||||||||||| ||| ||| |||||||||||||||||
Db 126 TGCTTCCACATGAAGACCTGGAAGCCCAGCCACTGTTTACCTTTTCAATTTGGCCGTGGCT 185

Qy 181 GATTCCCCTTATGATCTGCCTCGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 186 GATTCCCCTTATGATCTGCCTCGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 245

Qy 241 TGGGCTTGTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 246 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 305

Qy 301 GGGAGCACCGTGTGTCCTTACGGTGGTGGCTGCCGACAGGTATTCAAAGTGGTCCACCCC 360
             ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 306 GGGAGCACGTGTGTCCTTACGGTGGTGGCTGCCGACAGGTATTCAAAGTGGTCCACCCC 365

Qy 361 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGCGCTGGCATGTCTGCACCCTGTGG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 366 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGCGCTGGCATGTCTGCACCCTGTGG 425

Qy 421 GCCCTGGTCATCCGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 426 GCCCTGGTCATCCGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 485

Qy 481 ACGGCCGGTCCTCGTGAGAGCTTCATCATGAGGCTGGCCAATGCTGGCATGACATCATG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486 ACGGCCGGTCCTCGTGAGAGCTTCATCATGAGGCTGGCCAATGCTGGCATGACATCATG 545

Qy 541 TTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTTATTTTGCTCCCTCAAGATTGTT 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 546 TTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTTATTTTGCTCCCTCAAGATTGTT 605

Qy 601 TGGAGCTTGAGGCGGAGGCAGCTGGCCAGACAGCTGGATGAAGAGGCGACCCG 660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 606 TGGAGCTTGAGGCGGAGGCAGCTGGCCAGACAGCTGGATGAAGAGGCGACCCG 665

Qy 661 TTCATCATGGTGGTGCAATTGTGTTCATCACATGCTACCCTGCCCAGCGTGTCTGCTAGA 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 666 TTCATCATGGTGGTGCAATTGTGTTCATCACATGCTACCCTGCCCAGCGTGTCTGCTAGA 725

Qy 721 CTCTATTCCTCTGGACGGTGCCCTGCCCGAGTGCCTGCCGGATCCCCGTCCCATGGGGCCCTG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 726 CTCTATTCCTCTGGACGGTGCCCTGCCCGAGTGCCTGCCGGATCCCCGTCCCATGGGGCCCTG 785

Qy 781 CACATAACCCTGCAGCTTCACCTACATGAACAGACAGCCGGATCCCTGGTGTATTATTT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 786 CACATAACCCTGCAGCTTCACCTACATGAACAGACAGCCGGATCCCTGGTGTATTATTT 845

Qy 841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 846 TCAAGCCCCTCCTTCCCAAATTCTACACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG 905

Qy 901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 906 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 965

Qy 961 AGGAGTTGCCATCAGTGTGGCCAAATAGTTTCCAAAGCCCAGTCTGATGGCAATGGGATCCC 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 966 AGGAGTTGCCATCAGTGTGGCCAAATAGTTTCCAAAGCCCAGTCTGATGGCAATGGGATCCC 1025

Qy 1021 CACATTGTGAGTGGCAC 1038
              |||||||||||||||
Db 1026 CACATTGTGAGTGGCAC 1043

RESULT 14
ABA81529
ID ABA81529 standard; DNA; 1050 BP.
XX
AC ABA81529;
XX

```
CC    to determine the presence of or predisposition to a disease associated
CC    with altered levels of the polypeptide in mammals (especially humans) by
CC    detecting alterations in polypeptide expression levels relative to
CC    control samples. They are useful to identify agents binding polypeptide
CC    (e.g. cellular receptors or downstream effectors) and/or agents
CC    modulating cellular polypeptide expression or activity, useful as
CC    antagonists and agonists in disease treatment
XX
SQ    Sequence 1050 BP; 211 A; 299 C; 279 G; 261 T; 0 U; 0 Other;

Query Match      100.0%;   Score 1038;   DB 4;   Length 1050;
Best Local Similarity   100.0%;   Pred. No. 5e-287;
Matches 1038;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy         1 ATGTACAACGGGTCGTCGTGCTGCCCATGCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6 ATGTACAACGGGTCGTCGTGCTGCCCATGCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 65

Qy        61 CTGCTCATTGTGGCCTTTGTGCTGGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        66 CTGCTCATTGTGGCCTTTGTGCTGGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 125

Qy       121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTGGCCGTGGCT 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       126 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTTCAATTGGCCGTGGCT 185

Qy       181 GATTTCCTCCTATGATCGCCTGCCTTTCGACAGACTATTACCTCAGACGTAGACAC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       186 GATTTCCTCCTATGATCGCCTGCCTTTCGACAGACTATTACCTCAGACGTAGACAC 245

Qy       241 TGGGCTTTGGGGACATTCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       246 TGGGCTTTGGGGACATTCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGC 305

Qy       301 GGGAGCATGTGTGTTCCTTACGTGTGGCTGCCGGACACAGTATTCAAAGTGGTCCACCCC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       306 GGGAGCATGTGTGTTCCTTACGTGTGGCTGCCGGACACAGTATTCAAAGTGGTCCACCCC 365

Qy       361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGCTGGCCATCGTCTGCACCCTGTGG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       366 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGCTGGCCATCGTCTGCACCCTGTGG 425

Qy       421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCCGGTGCAAGAG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       426 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCCGGTGCAAGAG 485

Qy       481 ACGGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCATGGCTCGGCCATGACATCATG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       486 ACGGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCATGGCTCGGCCATGACATCATG 545

Qy       541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTTATTTGCTCCTCCAAGAATTGTT 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       546 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTTATTTGCTCCTCCAAGAATTGTT 605

Qy       601 TGGAGCCTGAGCGGGAGGCGGAGCAGCTCGGCAGCAGGAGGCGACCCCGG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       606 TGGAGCCTGAGCGGGAGGCGGAGCAGCTCGGCAGCAGGAGGCGACCCCGG 665

Qy       661 TTCATCATGGTGGTGCAATTGGTTCATCACATGCGTCTACCTGCCCAGCGTGTCTGCTAGA 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       666 TTCATCATGGTGGTGCAATTGGTTCATCACATGCGTCTACCTGCCCAGCGTGTCTGCTAGA 725

Qy       721 CTCTATTTGTGGACGGTGCCCTGAGTGGCCGGATCCCTGTCCATGGGGCCCTG 780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       726 CTCTATTTGTGGACGGTGCCCTGAGTGGCCGGATCCCTGTCCATGGGGCCCTG 785

Qy       781 CACATAACCCTGCAGCTTCACCTACATGAACAGCATGCTGGATCCCCTGTGTATTATTTT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       786 CACATAACCCTGCAGCTTCACCTACATGAACAGCATGCTGGATCCCCTGTGTATTATTTT 845

Qy       841 TCAAGCCCCTCCTTCCAAATTCTACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       846 TCAAGCCCCTCCTTCCAAATTCTACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG 905

Qy       901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       906 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 965

Qy       961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       966 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1025

Qy      1021 CACATTGTTGAGTGGCAC 1038
             ||||||||||||||||||
Db      1026 CACATTGTTGAGTGGCAC 1043


RESULT 15
ADE40271
ID   ADE40271 standard; cDNA; 1050 BP.
XX
AC   ADE40271;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human NOV41a cDNA - SEQ ID 177.
XX
KW   NOVx; cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;
KW   antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;
KW   antiparkinsonian; antiasthmatic; gynaecological; cardiomyopathy;
KW   atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;
KW   multiple sclerosis; graft-versus-host disease; Alzheimer's; Parkinson's;
KW   asthma; fertility disorder; vaccine; gene therapy; chromosome mapping;
KW   tissue typing; human; NOV; ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003064589-A2.
XX
PD   07-AUG-2003.
XX
PF   02-AUG-2002; 2002WO-US024483.
XX
PR   02-AUG-2001; 2001US-0309501P.
PR   03-AUG-2001; 2001US-0310291P.
PR   07-AUG-2001; 2001US-0310544P.
PR   08-AUG-2001; 2001US-0310951P.
PR   09-AUG-2001; 2001US-0311292P.
PR   13-AUG-2001; 2001US-0311979P.
PR   16-AUG-2001; 2001US-0312892P.
PR   17-AUG-2001; 2001US-0313201P.
PR   17-AUG-2001; 2001US-0313415P.
PR   20-AUG-2001; 2001US-0313643P.
PR   20-AUG-2001; 2001US-0313702P.
PR   21-AUG-2001; 2001US-0314031P.
PR   23-AUG-2001; 2001US-0314466P.
PR   28-AUG-2001; 2001US-0315403P.
PR   29-AUG-2001; 2001US-0315853P.
PR   17-SEP-2001; 2001US-0322716P.
PR   21-SEP-2001; 2001US-0323994P.
PR   14-DEC-2001; 2001US-0340233P.
PR   05-FEB-2002; 2002US-0354591P.
PR   19-MAR-2002; 2002US-0365478P.
PR   19-APR-2002; 2002US-0373814P.
PR   19-APR-2002; 2002US-0373825P.
PR   19-APR-2002; 2002US-0373989P.
PR   23-APR-2002; 2002US-0374632P.
PR   07-JUN-2002; 2002US-0386971P.
PR   01-AUG-2002; 2002US-00210172.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Kekuda R, Miller CE, Patturajan M, Pena CBA, Rieger DK;
PI   Shimkets RA, Zerhusen BD, Li L, Ji W, Padigaru M, Casman SJ;
PI   Voss EZ, Boldog FL, Gorman L, Leite MW, Vernet CAM, Anderson DW;
PI   Guo X, Zhong M, Gerlach VL, Hjalt T, Rastelli L, Spytek KA;
```

PI   Edinger SR., Elleman K., Malyankar UM, Macdougall JR, Stone DJ;
PI   Alsobrook JP, Lepley DM, Burgess CE, Majumder K, Wolenc AR;
PI   Smithson G;
XX
DR   WPI: 2003-663472/62.
DR   P-PSDB: ADE40272.
XX
PT   New NOVX polypeptides and nucleic acids, useful for preventing or
PT   treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT   atherosclerosis or diabetes, and in chromosome mapping, tissue typing or
PT   pharmacogenomics.
XX
PS   Claim 20; SEQ ID NO 177; 560pp; English.
XX
CC   The invention relates to a novel NOVX polypeptide. The polypeptide of the
CC   invention demonstrates cardiant, antiarteriosclerotic, hypotensive,
CC   cytostatic, anorectic, antidiabetic, immunosuppressive, anti-HIV,
CC   neuroprotective, nootropic, antiparkinsonian, antiasthmatic and
CC   gynaecological activities and may be useful in diagnosing, treating or
CC   preventing NOVX-associated disorders including cardiomyopathy,
CC   atherosclerosis, hypertension, cancer, obesity, diabetes, AIDS, multiple
CC   sclerosis, graft-versus-host disease, Alzheimer's disease, Parkinson's
CC   disease, asthma or fertility disorders. Furthermore, the polypeptides may
CC   be utilised as vaccines whilst the nucleic acids may be used as
CC   hybridisation probes, in gene therapy, chromosome mapping, tissue typing,
CC   preventive medicine and pharmacogenomics. The current sequence is that of
CC   the human NOV cDNA of the invention.
XX
SQ   Sequence 1050 BP; 211 A; 298 C; 279 G; 262 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1038; DB 9;   Length 1050;
Best Local Similarity   100.0%;   Pred. No. 5e-287;
Matches 1038;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGGTGATGCCGCCG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGGTGATGCCGCCG   65

Qy   61 CTGCTCATTGTGCCTTTGTGCTCGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 CTGCTCATTGTGCCTTTGTGCTCGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 125

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTACCTTTCAATTGGCCGTGGCT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTACCTTTCAATTGGCCGTGGCT 185

Qy   181 GATTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186 GATTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 245

Qy   241 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGGAACAGGGCC 305

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGACACTGTATTTCAAAGTGGTCCACCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGACACTGTATTTCAAAGTGGTCCACCCC 365

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCGTGGCATCGTCTGCACCCTGTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCGTGGCATCGTCTGCACCCTGTGG 425

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 485

Qy   481 ACGGACCGCGTCTCCCTGTGAGAGCTTCATCATGGAGTCGGCCAAAGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486 ACGGACCGCGTCTCCCTGTGAGAGCTTCATCATGGAGTCGGCCAAAGGCTGGCATGACATCATG 545

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT 605

Qy   601 TGGAGGCCTGAGCGGAGGCGGAGCAGCTCGGCCAGACAGCAGCAGCTCTCGGATGAAGAAGGCGACCCG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606 TGGAGGCCTGAGGCGGAGGCGGAGCAGCTCGGCCAGACAGCAGCAGCTCTCGGATGAAGAAGGCGACCCCG 665

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 725

Qy   721 CTCTATTTCCTCTGGACGGTGCCCTCGGATCCTCGTCCATGGGGCCCTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 CTCTATTTCCTCTGGACGGTGCCCTCGGATCCTCGTCCATGGGGCCCTG 785

Qy   781 CACATAACCCCTCAGCTTCACCTACATGAACACGCATGCTGGATCCCTGGTGTATTATTTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786 CACATAACCCCTCAGCTTCACCTACATGAACACGCATGCTGGATCCCTGGTGTATTATTTT 845

Qy   841 TCAAGCCCTCCTTCCAAATTCTACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846 TCAAGCCCTCCTTCCAAATTCTACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG 905

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGTCGC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 965

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1025

Qy 1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db 1026 CACATTGTTGAGTGGCAC 1043

Search completed: July 3, 2004, 06:34:38
Job time : 523 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    July  3, 2004, 06:08:16 ; Search time 103 Seconds
                                    (without alignments)
                                    5592.612 Million cell updates/sec

Title:          US-10-076-260-1
Perfect score:  1038
Sequence:       1 atgtacaacgggtcgtgctg.........cccacattgtgagtggcac 1038

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:  682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:  1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued_Patents NA:*
           1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
           2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
           3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
           4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
           5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
           6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	367.6	35.4	1164	4	US-09-170-496D-107	Sequence 107, App
2	367.6	35.4	2051	4	US-09-016-434-1259	Sequence 1259, Ap
3	367.6	35.4	2051	4	US-09-023-555-1212	Sequence 1212, Ap
4	366	35.3	1164	1	US-09-170-496D-221	Sequence 221, App
5	229.6	22.1	545	2	US-07-915-966C-1	Sequence 1, Appli
6	229.6	22.1	545	2	US-08-771-182-1	Sequence 1, Appli
7	229.6	22.1	545	2	US-08-853-194-1	Sequence 1, Appli
8	188.8	18.2	1594	2	US-08-955-713-1	Sequence 3, Appli
9	188	18.1	1435	2	US-08-955-713-3	Sequence 3, Appli
10	139.2	13.4	960	3	US-09-130-749-1	Sequence 1, Appli
11	139.2	13.4	960	3	US-09-130-749-1	Sequence 1, Appli
12	137.6	13.3	960	4	US-09-170-496D-195	Sequence 195, App
13	136	13.1	960	4	US-09-170-496D-59	Sequence 59, Appl
14	109	10.5	1020	4	US-09-170-496D-31	Sequence 31, Appl
15	109	10.5	1900	4	US-09-016-434-1484	Sequence 1484, Ap
16	109	10.5	1901	3	US-08-153-848-43	Sequence 43, Appl
17	109	10.5	1901	4	US-09-299-843A-43	Sequence 43, Appl
18	109	10.5	1901	4	US-09-088-337B-43	Sequence 43, Appl
19	109	10.5	2453	5	PCT-US93-11153-43	Sequence 43, Appl
20	109	10.5	1901	5	PCT-US95-07180-1	Sequence 1, Appli
21	105.8	10.2	1020	4	US-09-170-496D-181	Sequence 181, App
22	105.8	10.2	1098	4	US-09-170-496D-225	Sequence 225, App
23	105.2	10.0	1098	4	US-09-170-496D-117	Sequence 117, App
24	104.2	10.0	1597	2	US-08-724-974A-1	Sequence 1, Appli
25	104.2	10.0	1697	2	US-09-364-425B-26	Sequence 26, Appl
26	92.6	8.9	1842	1	US-08-442-134A-1	Sequence 1, Appli
27	92.6	8.9	1842	1	US-08-444-581B-1	Sequence 1, Appli
28	92.6	8.9	1842	1	US-08-446-088A-1	Sequence 1, Appli
29	92.6	8.9	2025	4	US-09-016-434-1482	Sequence 1482, Ap
30	89.8	8.7	3055	4	US-09-016-434-1456	Sequence 1456, Ap
31	82.4	7.9	1401	4	US-09-585-876-1	Sequence 1, Appli
32	81.4	7.8	775	4	US-08-148-708-8	Sequence 8, Appli
33	81.4	7.8	1179	4	US-08-148-708-3	Sequence 3, Appli
34	81.4	7.8	1378	1	US-08-759-848-2	Sequence 2, Appli
35	81.4	7.8	1378	4	US-08-148-708-6	Sequence 6, Appli
36	81.4	7.8	1378	4	US-09-016-434-1123	Sequence 1123, Ap
37	81.4	7.8	1378	5	PCT-US95-09383-2	Sequence 2, Appli
38	78.4	7.6	1155	3	US-09-053-866-3	Sequence 3, Appli
39	78.4	7.6	1155	4	US-09-479-130-3	Sequence 3, Appli
40	78.4	7.6	1155	4	US-09-472-130A-3	Sequence 3, Appli
41	78.2	7.5	1101	5	US-08-759-848-7	Sequence 7, Appli
42	78.2	7.5	1101	5	PCT-US95-09383-7	Sequence 7, Appli
43	78.2	7.5	3680	1	US-08-759-848-1	Sequence 1, Appli
44	78.2	7.5	3680	5	PCT-US95-09383-1	Sequence 1, Appli
45	75.8	7.3	1255	1	US-08-097-938-3	Sequence 3, Appli

```
                       ALIGNMENTS

RESULT 1
US-09-170-496D-107
; Sequence 107, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
;   APPLICANT: Behan, Dominic P.
;   APPLICANT: Chalmers, Derek T.
;   APPLICANT: Liaw, Chen W.
;   TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protei
;   TITLE OF INVENTION: Receptors
;   FILE REFERENCE: AREN-0040
;   CURRENT APPLICATION NUMBER: US/09/170,496D
;   CURRENT FILING DATE: 1998-10-13
;   NUMBER OF SEQ ID NOS: 294
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 107
;   LENGTH: 1164
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-170-496D-107

Query Match          35.4%;  Score 367.6;  DB 4;  Length 1164;
Best Local Similarity   63.9%;  Pred. No. 1e-96;
Matches  597;  Conservative   0;  Mismatches  319;  Indels  18;  Gaps   2;

Qy      1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCAGTGATGCGCCG   60
                                                            |||  ||||
Db     37 ATAGACAAGAAGAGACTGCTGTGTTCCGAGATGACTTCATTGCCAAGGTGTTGCCGCCG   96

Qy     61 CTGCTCATTGTGGCCTTTGTGCTGGGCGGCCACTAGGCAATGGGGGTCGCCCTGTGTGGTTTC   120
                                                    |||||
Db     97 GTGTTGGGGCTGCAGTTGGGGCTGGCATCCTTCTCGGCAATGGCCTGCCCTGTGGGATTTTC   156

Qy    121 TGCTTCCACATGAAGAGACCTGGAAGCCCAGCACACTGTTACCTTTTCAATTTGGCCGTGGCT   180
                                                    |||| ||   ||
Db    157 TGTTCCACCTCAAGTCCTGGAAATCCAGCCGGATTTTCCTGTTCAACCTGGCAGTAGCT   216

Qy    181 GATTCCTCCTTATGATCTGCCTGCCTCTTCGGACAGCATATTACCTCAGACGTAGACAC   240
                                                    |||  |||| ||
Db    217 GACTTTCCTACTGATCATCTGCCTCGCCGTTCGTGATGGACTACTATGTGCGGGCGTTCAGAC   276

Qy    241 TGGGCTTTTGGGGACATTCCCTCGCCAGCAGTGGGCTTCACGTTGGGCCATGAACAGGGCC   300
                                                    ||||
Db    277 TGGAACTTTGGGGACATCCCTTGCCATGATGTTGCCATGAACCTTCATGTTGCCACAGCCAG   336

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCCGACAGGTATTTCAAAGTGGTCCACCCC   360
                                                    ||   ||||||||||||
Db    337 GGCAGCATCATCTTCCTCACGGTGGGGTAGACAGGCATTTCCGGGTGGTCCATCCC   396

Qy    361 CACCACGCGGGTGAACACTATCTCCACCCGGGTGGCGGCGCTGGCATCGTCTGCACCCTGTGG   420
```

REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1259:
SEQUENCE CHARACTERISTICS:
LENGTH: 2051 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g219866
US-09-016-434-1259

Query Match   35.4%;   Score 367.6; DB 4;   Length 2051;
Best Local Similarity 63.9%;   Pred. No. 1.3e-96;
Matches 597; Conservative 0; Mismatches 319; Indels 18; Gaps 2;

```
Qy    1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACCACCATCTCCAGTGATGCGCGCG   60
        ||||| || |||   |||||| ||  |||||  ||  |  ||||  || |||  ||||||
Db   97 ATAGACAAGAAGAACTGCTGTGTTCCGAGATGACTTCATTGCCAAGAGTGTTGCCGCGG  156

Qy   61 CTGCTCATTGTGGCCTTTGTGTGGGGCCACAGGCAATGGGGTGGCCCGTGTGGTTTTC  120
        ||| |  || | |||  | ||  |||  |  ||||||| || || |||||| | ||||
Db  157 GTGTTGGGGCTGAGTTTATCTTTGGGCTTCTGGGCAATGGCCTTGCCCGTGGATTTC   216

Qy  121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  180
        || |||||| | |||  |||||||   |||| | ||||| ||| | |  |||| |||||
Db  217 TGTTTCCACCTCAAGTCCTGGAAATCCAGCCGGATTTTCCTGTTCAACCTGGCAGTAGCT  276

Qy  181 GATTTCCTTCCTATGATCTGCCCGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240
        ||  ||||| | |||||||||| |||| ||| | | ||  |  |||| || ||  |||||
Db  277 GACTTTCTACTGATCATCTGCCGCCGGTTCGTTGCGTGATGGACTACTATGTGCGGGTTCAGAC  336

Qy  241 TGGGCTTTGGGGACAATCCCTGCCGAGTGGGGGCTCTTGCACGTTGGCCATGAACAGAGGGCC  300
        |  |  |||||||| ||||||| ||||| | || ||| |||| || ||||| | ||||| ||||
Db  337 TGGAACTTTGGGGACATCCCTTGCCGGCGTGGTGCTCTTCATGTTTGCCATGAACCGGCCAG  396

Qy  301 GGGAGCATCGTGTCCTTACGGTGGTGGCCGGACAGGTATTCAAAGTGGTCCACCCC  360
        |||||||||||| ||||| ||||||||  || |||||| || |  |  | ||||| ||||
Db  397 GGCAGCATCATCTTCCTCACGGTGGTGGCGTAGACAGGTATTTTCCGGGTGGTGGTCCATCCC  456

Qy  361 CACCACGCGGTGAACACTATCTCCAATTGGACAGCAGCCATCATCTTCTGCCTTCTGTGTGG  420
        ||||||||||||||||||| ||| |||| |||||| ||| |||| |   ||| || ||||||
Db  457 CACCACGCCCTGAACCAGATCCCAATTGGACAGCAGCAGCCATCAGCCAAGATCAACCGCGG  516

Qy  421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGTGAGAACCATCTCTGCGTGTCAAGAG  480
        ||||||  | || ||| ||||||||  | |||  |||  | |||| |||||| || ||||||
Db  517 GGCATCACTGTTGGCCTAACAGTCCACCTCGAAGAAGAAGTTGCTGCTGATCCAGAATGGC  576

Qy  481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGAAGTGGCCAATGGCTGGCATGACATCATG  540
        | |||||||||  |||| ||||| |||| ||||||||| ||||| || || |  | |||
Db  577 CCTGCAAATGTGTCATCAGCTTCAGCCATCGCCATACCTTCCGGTGGCACGAAGCTATG  636

Qy  541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCCAGGAGGCAGCAGTCTCGGTGCCGATGAAGAAGGCGACCCGG  600
        |||||||||||||  | |||  || ||||||| | | |||||   ||||| || ||||| |||| ||||||
Db  637 TTCCTCCGGAGTTCTTCCTGCGGCCGCAGCAGCTGCTGCTGATCCAGAATTATTTTCGGGGTGGTGGTCACCCTGTG  696

Qy  601 TTGGAGGCCTGAGGGCGGAGGCAGCAGGTGGGCAAGAAGGCGACCCGG  660
        |||| |||||| | ||  | |||||| ||  ||| | ||  || ||||
Db  697 TGGAGCCTGAGGCGGAGGCAGCAGGAGG---ACAAATGACACGGCCATCACC  753

Qy  661 TTCATCATGGTGGTGGCAATGGTGCAATTGTGTTCATCACATTTCCGGTGGCAAGTTGTGTGCCTGCTGTAGA  720
        |||||||||| ||  | || || |||||||||||| |||  |||||| |||||  |   | | ||||||| |||
Db  754 TTCATCATGGTGGTGGCCATCGTCTTTGTCATCTGCTTTGTCATCGTCTTCCTCTTCCGGTTGCGTCG  813

Qy  721 CTCTATTTCCTCTGACGGTGCCCTC----------GAGTGCCTGCCGATCCCTCT  765
        ||  |   |||  | | | ||||||          || |||||| || | | ||
Db  814 ATCCGCCATCTTCTGGCTCCCTGCACACTTCGGGCACGCAGAATTGTGAAGTGTACCGCTCG  873

Qy  766 GTCCATGGGGCCCTGCACATAACCCTGCACTTCACCTCAGTTCTGTGTGCCTGC  885
        |  || ||||||| ||| |||| || | |||| |||  ||| | | | || || ||||||
Db  ...
```

Db   874 GTGGACCTGGCGTTCTTATCACTCTCAGCTTCACCTACATGAACAGCATGCTGGACCCC 933
          ||||||||||| || || ||||||||||| ||||||||||||| || ||||||
Qy   826 CTGGTGTATTATTTTTCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGC 885
          |||||| |||| ||||||||||| |||| ||||||||||||| |||| ||||||
Db   934 GTGGTGTACTACTTCTCCAGCCCATCCTTCTCCCAACTTCTTCCACTTTGATCAACCGC 993

Qy   886 AGCTGAAACCCAAGCAGCCAGGACACTCAAAAA 919
          |||||||||| ||| ||| |||||
Db   994 TGCCTCCAGAGGAAGGATGACAGGTGAGCCAGATA 1027

RESULT 3
US-09-023-655-1212
; Sequence 1212, Application US/09023655
Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1212:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2051 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g219866
US-09-023-655-1212

Query Match   35.4%; Score 367.6; DB 4; Length 2051;
Best Local Similarity 63.9%; Pred. No. 1.3e-96;
Matches 597; Conservative 0; Mismatches 319; Indels 18; Gaps 2;

Qy   1 ATGTACAACGGTCGTGCTCGCCGCATCGAGGGGACCACCATCTCCCAGGTGATGCCGCCG 60
          |||||||||||||||||||||||| |||||| ||||||||||||||||||
Db   97 ATAGACAAGAAGAACTGCGTGTGTTCCGAGATGACTTCATTGCCCAAGGTGTTGCCGCCG 156

Qy   61 CTGCTCATTGTGGCCTTGTGCTGGGCGCCACTAGGCAAATGGGGTCGCCCTGTGTGGTTTC 120
          |||||||||||||||||| |||| |||| ||| ||||||| |||||||
Db   157 GTGTTGGGCTGAGTTTATCTTTGGGCTTCTGCGCAATGGCCTTGCCCTGTGTGGGATTTTC 216

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180

Db   217 TGTTTCACCTCAAGTCCTGGAAATCCAGCCGGATTTTCCTGTTCAACCTGGCAGTAGCT 276
          |||||||| ||| || |||||||||||| ||||||||| ||||||
Qy   181 GATTCCTCCTATGATCTGGCCTGCCTTTCGGACAGACTATTACCTCAGACGAGTAGACAC 240
          ||||||| |||||||||||||| |||||||||| |||||||||||||||||
Db   277 GACTTTCTACTGATCATCTGCCTGCCGTTCGTGATGACTACTACGTGGCGGGTTCAGAC 336

Qy   241 TGGGCTTTGGGGACATCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGACAGGGCC 300
          ||||||||||||||||||| |||||||||| ||||||| ||||||||||
Db   337 TGGAACTTTGGGGACATCCCTGCCGGTGTCTCTCATGTTTGCCATGAACCGCCAG 396

Qy   301 GGGAGCATCGTCGTCTCCTACGGTGGTGCGTCGCGGACAGTATTCAAAGTGGTCTCACCCC 360
          ||||||||||||||| ||||||||||||||| ||||||||| |||||||||||
Db   397 GGCAGGATCATCTTCCTCACGTGGTGGCGGTAGACAGGTATTTCCGGGTGGTCCATCCC 456

Qy   361 CACCACGCGGTGAACACTATCTCACCCGGTGCGGCTGGCATCGTCTGCACCCGTGTG 420
          ||||||||| |||||||||||||||||||||| ||||| ||||| |||||||||
Db   457 CACCAGCCCCTGAACAGATCTCCAATTGGACAGCAGCCATCATCTCTGCCTTCTGTGTG 516

Qy   421 GCCCTGGTCATCCTGGAACAGTGTATCTTTTGCTGAGAACCATCTCTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||| ||||||||||||||||||
Db   517 GGCATCACTGTTGGCCTAACAGTCCACCTCCTGAAGAGAGAGTTGCTGATCCAGAATGGC 576

Qy   481 ACGGGCCGTCTCCTGTGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACGATCATG 540
          ||||| ||||||||| ||||||||||||||||||||||| |||||||||||
Db   577 CCTGCAAATGTGCATCAGCTTCAGCATCTGCCATCACCTTCCGGTGCGGCACGAAGCTATG 636

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTCCGCCATCATCTTATTTGCTCCTTCAAGATTGTT 600
          ||||||||| ||||| || |||||| |||||||||||||||| |||||||||
Db   637 TTCCTCCTGGAGTTCTTCCTGCCCCTGGGCATCATCCTGTTCTGTGCTCCAGATTATC 696

Qy   601 TGGAGCCTGAGCGGGAGGCAGCAGTCGCCAGCAGCACAGGCTCGGATGAAGAAGGGACCCG 660
          ||||||||| ||||||| |||||||||||||||||||||||||
Db   697 TGGAGCCTGCGCGGAGAG--ACAAATGACCGGCATGCCAAGATCAAGAGAGCCATCACC 753

Qy   661 TTCATCATGGTGGTGGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
          |||||||||||||||| ||||||||||||||||||| | || |||||||||
Db   754 TTCATCATGGTGGTGGGCCATCGTCTTGTCATCGTCTTCCTTCCTCAGCGTGGTGTGCGG 813

Qy   721 CTCTATTTCCTCTGGACGGTGCCCTC------------GAGTGCCTGCGATCCCTCT 765
          ||||||| ||| ||||||||||| | | |||||||| |||||
Db   814 ATCCGACATCTTCTGGCGTCCTGCCACACTTCGGSCACGCAGAATTGTGAAGTGTACCGCTCG 873

Qy   766 GTCCATGGGGGCCCTGCACATAACCCTCAGCTTCACTACATGGAACAGCATGCTGGATCCC 825
          |||||||||||||||||||||||||||||||| |||| ||||||||||||
Db   874 GTGGACCTGGCGTTCTTTATCACTCTCAGCTTCACTACATGAACAGCATGCTGGACCCC 933

Qy   826 CTGGTGTATTATTTTTCAAGCCCCTCCTTTCCCAAATTCTACACAACAAGCTCAAAATCTGC 885
          |||||||||||||||| ||||||||| | |||||||||||||||||||
Db   934 GTGGTGTACTACTTCTCCAGCCCATCCTTCTCCCAACTTCTTCCACTTTGATCAACCGC 993

Qy   886 AGTCTGAAACCCAAGCAGCCAGGACACTCAAAAA 919
          | | |||||||||| ||| |||||
Db   994 TGCCTCCAGAGGAAGGATGACAGGTGAGCCAGATA 1027

RESULT 4
US-09-170-496D-221
; Sequence 221, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 221
; LENGTH: 1164

```
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-170-496D-221

Query Match      35.3%;   Score 366; DB 4;  Length 1164;
Best Local Similarity  63.8%;  Pred. No. 2.9e-96;
Matches 596;   Conservative   0;  Mismatches 320;  Indels 18;  Gaps 2;

Qy      1  ATGTACAACGGGTCGTCGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG      60
Db     37  ATAGACAAGAAGAACTGCTGTGGTCGTTCCGAGATGACTTCATTGCCAAGGTGTTGCCGCCGG   96

Qy     61  CTGCTCATTGTGGGCCTTTGTCTGGGCGCACTAGGCAATGGGTCGCCCTGTGTGGTTTC     120
Db     97  GTGTTGGGGCTGGAGTTATCTTGGGCTTCTGGGCAATGGCCTTGCCCCTGTGGATTTTC    156

Qy    121  TGCTTCCACATGAGAACCTGGAAGCCAGCACTGTTTACCTTTTCAATTGGCCGTTGGCT    180
Db    157  TGTTTCCACCCTCAAGTCCTGGAAATCAGCGGATTTTCCTGTTCAACCTGGCAGTAGCT    216

Qy    181  GATTCCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC    240
Db    217  GACTTTCTACTGATCATCTGCCTGCCGTTCGTGATGACTACTATGTGCGGCGTTCAGAC    276

Qy    241  TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC    300
Db    277  TGGAAGTTTGGGGACATCCCTGCCGGCTGTGTGCCTCATGTTTGCCATGAACCGCCCAG    336

Qy    301  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC    360
Db    337  GGCAGCATCATCTTCCTCACGGTGGTGGGCCGTAGACAGGTATTTCCGGGTGGTCCATCCC    396

Qy    361  CACCACGCGGTGAACACTATCTCCACCCGGGTGGGCGTCTGCACCCGTGTGG    420
Db    397  CACCACGCCCTGAACAGAGATCTCCAATTGACAGCAGCCATCATCTCTTGCCTTCTGTGG    456

Qy    421  GCCCTGGTCATCCTGGGACAGTGTATCTTTTGCTCGAGAACCATCTCTGCGGTGCAAGAG    480
Db    457  GGCATCACTGTGGCCTAACAGTCCACCTCCTGAAGAAGAAGTTGCTGATCCAGAATGGC    516

Qy    481  ACGGCCGTCTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGTGGCATGACATCATG    540
Db    517  CCTGCAAATGTGTGCATCAGCTTCAGCATCTGCAGCATCTGCCATACCTTCCGGTGGCACGAAGCTATG    576

Qy    541  TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCCCTTCAAGATTGTT    600
Db    577  TTCCTCCTGGAGTTCCTTGCTCCTGCCCCTCGGCATCATCTTCTGTCTCGCTCAGCCAGAATTATC    636

Qy    601  TGGAGGCCTGGAGCGGAGGCCAGCAGCTGCCCAGACAGGCTCGGATGAAGAAGACGGCACCCGG    660
Db    637  TGGAGGCCTGCGGCAGAG--ACAAATGGACCCGGCATGCCAAGATCAAGAGAGCCAAAACC    693

Qy    661  TTCATCATGGTGGTGGCAATTGTGTTCATCACCATGCTACCTGCCCAGCCGTGTCTGCTAGA    720
Db    694  TTCATCATGGTGGTGGCCATCGTCTTGTCATCGTCGTCTGTCTTCATCACCGTGAACCGTGGCGCG    753

Qy    721  CTCTATTCCTCCTGAQCGGTGCCCTC------------GAGTGCTTGCGATCCTCT    765
Db    754  ATCCGGACATCTTCTGTCCTCCTGCACTTCGGGCACGCAGAATTGTGAAGTGTACCGCTCG    813

Qy    766  GTCCATGGGGCCCTGCACATAACCCTCAGCTTCACCACCACATGAACATGACTGATCCC    825
Db    814  GTGGACCTGCGGTTCTTTATCACTTCTCAGCTTCACTCAGTCCCACATGCTACAGCTGATGACCCC    873

Qy    826  CTGGTGTATTATTTTCAAGCCCCTCCTTTCACAACAATTCACACAGAAGTCTAAAAATCTGC    885
Db    874  GTGGTGTACTTCTTCCAGGCCACTTGTCTTCTTCCTCCACTTTGATCAACCGC    933

Qy    886  AGTCTGAAACCCAGCAGCCAGGAGACACTCAAAA    919
Db    934  TGCCTCCAGAGGAAGGATGACAGGTGAGCGAGATA    967
```

RESULT 5
US-07-915-966C-1
; Sequence 1, Application US/07915966C
; Patent No. 5668006
; GENERAL INFORMATION:
;   APPLICANT: Hadcock Dr., John R.
;   APPLICANT: Ozenberger Dr., Bradley A.
;   APPLICANT: Pausch Dr., Mark H.
;   TITLE OF INVENTION: Receptor Identification Method
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: American Home Products Corporation
;   STREET: One Campus Drive
;   CITY: Parsippany
;   STATE: New Jersey
;   COUNTRY: USA
;   ZIP: 07054
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/915,966C
;   FILING DATE: 17-JUL-1992
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Matthews, Gale M.
;   REGISTRATION NUMBER: 32,269
;   REFERENCE/DOCKET NUMBER: 31,829-00
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 201-683-2134
;   TELEFAX: 201-683-4117
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 545 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;   ORGANISM: Rat
;   PUBLICATION INFORMATION:
;   AUTHORS: Hadcock Dr., John R.
;   AUTHORS: Dr. Ozenberger, Bradley A.
;   AUTHORS: Dr. Pausch, Mark H.
;   TITLE: Receptor Identification Method
;   DATE: 17-JUL-1992
US-07-915-966C-1

Query Match   22.1%;   Score 229.6; DB 1; Length 545;
Best Local Similarity 65.0%; Pred. No. 7.3e-57;
Matches 356;   Conservative   0; Mismatches 189; Indels 3; Gaps 1;

```
Qy    157  TACCTTTTCAATTGGCCGTTGGCTGATTCCCTCCTTATGATCTGCCTGCCTTTTCGGACA    216
Db      1  TTCGTGTGACCCTGGTCGGGGCGCTGACTTTCTGCTTGCCGTTCTGCCGTTCTGACG      60

Qy    217  GACTATTACCTCAGACGTAGACACTGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTC    276
Db     61  GACAACTATGTCCAGAACTGGGACTGGAAGTTCGGGACGATCCCCCCTGGCGCGCGATGCTC    120

Qy    277  TTCACGTTGGCCATGAACAGGGCCGGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGAC    336
Db    121  TTCATGTTGGCCATGAACAGCGACACAGGGCAGCATCATCTTCCTCACGCACGCGTGGTGTGGAC    180

Qy    337  AGTTATTTCAAAGTGGTCCACCCCCACCACGCGGTGAACACTATCTCCACCCGGGTGGCG    396
Db    181  AGTACTTCAGGGTGGTCCACCCTGCCACCACTTCCTGAACAGAATCTCCAACCGGACCGCG    240

Qy    397  GCTGGCATCGTGCTGCACCCTGTGGGCCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTG    456
```

; DATE: 20-DEC-1996
US-08-771-182-1

Query Match   22.1%; Score 229.6; DB 2; Length 545;
Best Local Similarity 65.0%; Pred. No. 7.3e-57;
Matches 356; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

```
Qy  157 TACCTTTTCAATTGGCCGTGGCTGATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACA 216
            ||||   |||    ||||||||||| |  |||  ||||  | |||| || ||  ||||
Db    1 TTCGTGGTGAACCTGGTCGGGGTGACTTTCCTCGTCATTTGCTTGCGTTCTTGACG   60

Qy  217 GACTATTACCTCAGAGCTAGACACTGGGCTTTTGGGGACAATTCCCTGCCGAGTGGGGCTC 276
            ||||| |||   ||||||||||||  |  ||||   ||| ||||||||||||||||
Db   61 GACAACTATGTCCAGAACTGGACTGGAGGTTCGGGAGCATCCCCTGCCCGCGTGATGCTC  120

Qy  277 TTCACGTTGGCCATGAACAGAGGCCGGAGCATCGTGTTCCTTACGTGGTGGTCGCGGAC 336
            ||||||||||||||||||  ||| |||| ||| ||||||||| ||||  ||  | |||
Db  121 TTCATGTTGGCCATGAACCGAGCCGCAGCATCATCTTCCTCACGTGGTGGTCGTGTGAC  180

Qy  337 AGGTATTTCAAAGTGGTCACCCCACCACGGTGAACACTATCTCACCCGGTGGCG 396
            |||| |||    |    |||   |   ||||  ||||||| ||||||  ||||
Db  181 AGGTACTTCAGGGTGGTGGTCCACCCGCCACTTCCTGAACAAGATCTCAACCGGACGGCG  240

Qy  397 GCTGGCATCGTCTGCACCCTGTGGGCCCTGGTCATCCTGGGAACAGTGTACTTTTGCTG 456
            ||||||||||| ||||  || || ||| |||||||||||||   ||  |  | ||||||
Db  241 GCCATCATCTTCTGCTTCTGTGGGGCATCACCATCGGCCTGACAGTCCACCTCCTCTAC  300

Qy  457 GAGAACCATCTCTGGTGCAAGAGACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCG 516
            ||||||||||| ||||  ||||  || || | ||||| |||||||||||||||| ||
Db  301 ACGGACATGATGACCCGAAACGGCGAATGCAAAACCTGTGCAGCAGTTTTAGCATCTGCTAC  360

Qy  517 GCCAATGGCTGGCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATC 576
            ||||||||||||||||| || |||||||| || |||||||  ||||||  ||||||||||
Db  361 ACTTTCAGGTGGCACGATGCAATGTTCCTTGGAATTCTTCCTGCCCTGGGCATCATC  420

Qy  577 TTATTTTGCTCCTTCAAGATTGTTTGGAGCCTGAGGCGGAGGCAGCGGCCAGACAG 636
            |      ||| |||||||||| |||||||||||| | |||  |||| ||||| | ||
Db  421 CTGTTCTGCTCTGGCAGGATCATTTGGAGCCTAAGGCAGAG--ACAGATGGACAGGCAC  477

Qy  637 GCTCGGATGAAGAAGGCGACCCGGTTCATCATGGTGGTGGCAATTGTGTTCATCACATGC 696
            ||||||||| | ||||| || |||||||||||||| |||||||||||||  | ||||||
Db  478 GTCAAGATCAAGAGAGGGCCATCAACTTCATCATGTGGTTGCCATTGTGTTTGCCATCTGC  537

Qy  697 TACCTGCC 704
            ||||||
Db  538 TGGCTGCC  545
```


RESULT 6
US-08-771-182-1
; Sequence 1, Application US/08771182
; Patent No. 5929209
; GENERAL INFORMATION:
;,   APPLICANT: Hadcock Dr., John R.
;,   APPLICANT: Ozenberger Dr., Bradley A.
;,   APPLICANT: Pausch Dr., Mark H.
;,   TITLE OF INVENTION: Receptor Identification Method
;,   NUMBER OF SEQUENCES: 19
;,   CORRESPONDENCE ADDRESS:
;,   ADDRESSEE: American Home Products Corporation
;,   STREET: One Campus Drive
;,   CITY: Parsippany
;,   STATE: New Jersey
;,   COUNTRY: USA
;,   ZIP: 07054
;,   COMPUTER READABLE FORM:
;,   MEDIUM TYPE: Floppy disk
;,   COMPUTER: IBM PC compatible
;,   OPERATING SYSTEM: PC-DOS/MS-DOS
;,   SOFTWARE: PatentIn Release #1.0, Version #1.30
;,   CURRENT APPLICATION DATA:
;,   APPLICATION NUMBER: US/08/771,182
;,   FILING DATE: 20-DEC-1996
;,   ATTORNEY/AGENT INFORMATION:
;,   NAME: Matthews, Gale F.
;,   REGISTRATION NUMBER: 32,269
;,   REFERENCE/DOCKET NUMBER: 31,829-D1
;,   TELECOMMUNICATION INFORMATION:
;,   TELEPHONE: 201-683-2134
;,   TELEFAX: 201-683-4117
;, INFORMATION FOR SEQ ID NO: 1:
;,   SEQUENCE CHARACTERISTICS:
;,   LENGTH: 435 base pairs
;,   TYPE: nucleic acid
;,   STRANDEDNESS: single
;,   TOPOLOGY: linear
;,   MOLECULE TYPE: cDNA
;,   HYPOTHETICAL: NO
;,   ANTI-SENSE: NO
;,   ORIGINAL SOURCE:
;,   ORGANISM: Rat
;,   PUBLICATION INFORMATION:
;,   AUTHORS: Hadcock Dr., John R.
;,   AUTHORS: Dr. Ozenberger, Bradley A.
;,   AUTHORS: Dr. Pausch, Mark H.
;,   TITLE: Receptor Identification Method

RESULT 7
US-08-853-194-1
; Sequence 1, Application US/08853194
; Patent No. 6077666
; GENERAL INFORMATION:
;,   APPLICANT: Hadcock Dr., John R.
;,   APPLICANT: Ozenberger Dr., Bradley A.
;,   APPLICANT: Pausch Dr., Mark H.
;,   TITLE OF INVENTION: Receptor Identification Method
;,   NUMBER OF SEQUENCES: 19
;,   CORRESPONDENCE ADDRESS:
;,   ADDRESSEE: American Cyanamid Company
;,   STREET: One Cyanamid Plaza
;,   CITY: Wayne
;,   STATE: NJ
;,   COUNTRY: United States of America
;,   ZIP: 06904-0060
;,   COMPUTER READABLE FORM:
;,   MEDIUM TYPE: Floppy disk
;,   COMPUTER: IBM PC compatible
;,   OPERATING SYSTEM: PC-DOS/MS-DOS
;,   SOFTWARE: PatentIn Release #1.0, Version #1.25
;,   CURRENT APPLICATION DATA:
;,   APPLICATION NUMBER: US/08/853,194
;,   FILING DATE:

```
   CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
   APPLICATION NUMBER: US 07/915,966
   FILING DATE: 17-JUL-1992
 ATTORNEY/AGENT INFORMATION:
   NAME: Tsevdos Dr., Estelle J.
   REGISTRATION NUMBER: 31,145
   REFERENCE/DOCKET NUMBER: 31829-00
 TELECOMMUNICATION INFORMATION:
   TELEPHONE: 203-321-2361
   TELEFAX: 203-321-2971
   TELEX: 710-474-4059
 INFORMATION FOR SEQ ID NO: 1:
   SEQUENCE CHARACTERISTICS:
     LENGTH: 545 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: single
     TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   HYPOTHETICAL: NO
   ANTI-SENSE: NO
   ORIGINAL SOURCE:
     ORGANISM: Rat
   PUBLICATION INFORMATION:
     AUTHORS: Hadcock Dr., John R.
     AUTHORS: Dr. Ozenberger, Bradley A.
     AUTHORS: Dr. Pausch, Mark
     TITLE: Receptor Identification Method
     DATE: 17-JUL-1992
 US-08-853-194-1

 Query Match    22.1%;  Score 229.6;  DB 3;  Length 545;
 Best Local Similarity  65.0%;  Pred. No. 7.3e-57;
 Matches  356;  Conservative  0;  Mismatches 189;  Indels  3;  Gaps  1;

 Qy   157 TACCTTTTCAATTGGCCGTGGCTGGTGATTTCCTCCTTATGATCTGGCCTGCCTTTTCGGACA 216
              ||||||||||||| |||||||||  ||||||
 Db     1 TTCGTGGTGAACCTGGTCGGGGGCTGACTTTCTCCTGATCATTGCTTGCCGTTCTTGACG 60

 Qy   217 GACTATTACCTCAGACGTAGACACTGGGCTTTGGGACATCCCTGCCGAGTGGGGCTTC 276
 Db    61 GACAACATGTCCAGAACTGGGACTGGAGGTTCGGGAGCATCCCTGCGCGCGTGATGCTC 120

 Qy   277 TTCAACGTTGGCCATGAACAGAGGCCGGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGAC 336
 Db   121 TTCATGTTGGCCATGAACGACAGAGGGCAGCATCATCTTCCTCACCGTGGTGGCTGTGGAC 180

 Qy   337 AGGTATTTCAAAGTGGTCACCCCACCACACGCGGTGAAACACTATCTCCACCGGGTGGCG 396
 Db   181 AGGTACTTCAGGGTGGTGGTCCACCCGGTCACCCTGGAACAAGATCCCAACCGGACGGGCG 240

 Qy   397 GCTGGGCATCGTCTGCACCCTGGTCACCCTGGGAACCAGTGTATCTTTTGCTG 456
 Db   241 GCCATCATCCTTGCTTGCTGTGGGCATCACCATCGGCCTGACAGTTCCACCTCCTCTAC 300

 Qy   457 GAGAACCATCCTGCGTGCAAGAGACGGCCGTCTCCTGTGAGCTTCATCATGGAGTCG 516
 Db   301 ACGGACAATGATGACCCGAAACGCGATGCAAACCTGTCAGCAGTTTAGCATCGCTAC 360

 Qy   517 GCCAATGGCTGGCATGAACATCATGTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATC 576
 Db   361 ACTTTCAGGTGGCACGATGGCGCAATGTTCCTCCTTGGAATTCTTCCTGCCCCCCTGGGGCATCATC 420

 Qy   577 TTATTTTGCTCCTTCAAGATTGTTTGGAGGCCTGAGGCCGGAGGCGAGCAGTGGCCAGACAG 636
 Db   421 CTGTTCTGCTCTGCCAGGATCATTTGGAGGCCTAAGGCAGGCAGAG--ACAGATGGAGACAGGCAC 477

 Qy   637 GCTCGGATGAAGAAGAGGCCGACCCGGTTCATCATGGTGGTGCAATTGTGTTCATCACATGC 696
 Db   478 GTCAAGATCAAGAGAGGGCCATCAACTTCATCATGTGGTTGCCATTGTGTTGCCATCTGC 537

 Qy   697 TACCTGCC 704
              ||||||
 Db   538 TGGCTGCC 545
```

```
 Query Match    18.2%;  Score 188.8;  DB 2;  Length 1594;
 Best Local Similarity  54.7%;  Pred. No. 8.6e-45;
 Matches  445;  Conservative  0;  Mismatches 357;  Indels  12;  Gaps  3;

 Qy    53 TGCCCGCCGTGCTCATTGTGGCCTTGTGCTGGGCCCACTAGGCAATGGGGTCGCCCTGT 112
 Db   519 TGGCACCAATCCTGGCCCTGGAGTTTGTCCTGGCCTGGTGGGGAACAGTTTGGCCCTCT 578

 Qy   113 GTGGTTTCTGCTTCCACATGAAGACCTGGAAGCCCAGCACCCACACTGTTTACCTTTTCAATTTGG 172
 Db   579 TCATCTTCTGCATCCATCACCGCGGCCCTGGACCTCCAACACCGGTGTTCCTGGTCAGCCTGG 638

 Qy   173 CCGTGCTGATTTCCCTTATGATCTGGCCTTTGCCTGCCTCCCTCCGCCTGGACCTGGCCAGACGAC 232
 Db   639 TGGCCCTGACTTCCTCCTGATCAGCAACCTGCCCCTCCCGCCTGGACTACTACCTCCTCC 698

 Qy   233 GTAGACACTGGGCTTTGGGACATTCCTGCGGTGGGCTCTTCACGTTGGCCATGA 292
 Db   699 ATGAGACCTGGCGCTTTGGGCGGCTGCTGCTGCCTGAAAGTCAACCTCTTTCATGCGTCCACCA 758

 Qy   293 ACAGGGCCGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGTGGTATTTCAAAGTGG 352
 Db   759 ACGGCACGGCCAGGGTTGTCTTCCTCCTCACAGCCATCCGCACTCAACCCGCTACCTGAAGGTGG 818
```

TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1435 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-955-713-3

Query Match   18.1%;   Score 188;   DB 2;   Length 1435;
Best Local Similarity 54.1%;   Pred. No. 1.4e-44;
Matches 440;   Conservative 3;   Mismatches 359;   Indels 12;   Gaps 3;

```
Qy     53 TGCCGCCGCTCCATTGTGGCCTTTGTGCTGGCCGCACTAGGCAATGGGGTCGCCCTGT 112
           |||| |||  |||||  ||||||||  |||| |||||||
Db     18 TGGCACCAATCCTGGCCCTGGAGTTTGTCCTGGGCCTGGTGGGGAACAGTTTGGCCCTCT 77

Qy    113 GTGGTTTCTGCTTCCACATGAAGACCTGGAAGCCCAGGCAGACTGTTTACCTTTCAATTTGG 172
           ||||| |  |||| |||  |||  ||||| ||| | |||||
Db     78 TCATCTTCTGCATCCACACGCGGCCCCTGGACCTCCAACACGGTGTTCCTGGTCAGCCTGG 137

Qy    173 CCGTGGCTGATTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACATTACCTCAGAC 232
           |||| |||  |||| |||  || ||  |||||  ||||
Db    138 TGGCCGCTGACTTCCTCCCTGATCAGCAGACAACCTGCCCCTCCGGGACTACTACCTCCTCC 197

Qy    233 GTAGACACTGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGA 292
           |||||||||  |||| |||||| |||  ||| |||
Db    198 ATGAGACCTGGCGCTTTGGGCGTGCTGCCTGCCAAAGTCAACCTTCTTCATGCTGTCNACCA 257

Qy    293 ACAGGGCGCGGGAGCATCGTGTTCCTTACGGTGGTGCGGACAGGTATTCAAAGTTGG 352
           |||| ||  |||||  ||| |||||||  ||| ||
Db    258 ACCGCAAGGCCAGCGTTGTCTTCCTCACAGCCATCAACCGCTCACCTGAAGGTGG 317

Qy    353 TCCACCCCCACCACGCGGCGGTGAACACTATCTCCACCCGGTGGGCGGCACNCTGCCCGGGTGGNCGGGG 412
           || |||||||||||| ||||||| |  || |||
Db    318 TGCANCCCCACCACGGTGCGAACCGTGCTCTTCCGTGGGGAACCATCCTCGG 377

Qy    413 CCCTGTGGGCCCTGGCTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCCTCGG 472
           ||||| |||||||  ||||||||| |
Db    378 GAATCTGG---TGGGCATCCTGCTCCTCAACGGNACCTGCTCCTGGAACACCTTCTCCG 434

Qy    473 TGCAAGAGACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGCTGGCCAATGGCTGGCATG 532
           |||  |||| | ||||  || ||  ||||
Db    435 GCCCCCTCCTGCCTCAGCTACAGGGTGGGCACGAARCCCTCGGCCTCCGCTGGCACC 494

Qy    533 ACATCATGTTCCAGCTGGAGTTCTTTATGCCCCTCGCCATCATCTATTTTGCTCCTCA 592
           || |||||| |||||||| | ||  |||  || | |||| ||
Db    495 AGGCACTGTACCTGCTGGARTYTTCTGCCACTGGCCCTCATCCTCTTT--GCTATTG 551

Qy    593 AGATTGTTTGGAGCCTGAGCGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGG 652
           ||  |||| |||||  ||||||
Db    552 TGAGCATTGGGCTCACCATCCGGAACCGTGTCTCGGCGGGCAGGCAGCCCGGAGAGGGG 611

Qy    653 CGACCCGGTTCATCATGGTGGGCCAATTGTGTCATCACATGGCCCAGCCGTGT 712
           |||| |||| |||||||||| ||
Db    612 CCATGCGTGTNCTGGCCCATGGTGGTGCGYGTCACCACCATCTGCTTCTTGCCCCAGCATCA 671

Qy    713 CTGCTAGACTCTATTTCCTCTGACGGTGCCCTGCCGATCCCTGTCCATG 772
           ||||||| |||||  ||||||
Db    672 TCTTTGGCATGGCTTCCATGGGGTGGCCTTTCTGCGTCCGCCGTGCGGATCCCTCT 731

Qy    773 GGGCCCTGCACATAA------CCCTCAGCTTCACCTACACATGAACAGCATGCTGGATCCCC 826
           | |  ||| ||||  |||||||||||| | ||||||
Db    732 GCACACAGCTCTTCCATGGCCTTCCATGGCTTCACCTACCTCAACACAGTGTCCTGGACCCCG 791

Qy    827 TGGTGTATTATTTTCAAGCACCCCTCCTTTCCCAA 860
           ||| || |||||| ||||||  ||||
Db    792 TGCTCTACTGCTTCTTCTAGCCCAACTTCCTCCA 825
```

RESULT 9
US-08-955-713-3

Sequence 3, Application US/08955713
Patent No. 5955308
GENERAL INFORMATION:
APPLICANT: SATHE, GANESH
APPLICANT: MOONEY, JEFFREY
APPLICANT: BERGSMA, DERK
APPLICANT: HALSEY, WENDY
TITLE OF INVENTION: cDNA CLONE HEOAD54 THAT ENCODES   A HUMAN 7-TRANS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,713
FILING DATE: 23-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/050,124
FILING DATE: 18-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70087
TELECOMMUNICATION INFORMATION:

```
RESULT 10
US-09-130-749-1
; Sequence 1, Application US/09130749
; Patent No. 6031090
; GENERAL INFORMATION:
;   APPLICANT: SHABON, USMAN
;   APPLICANT: ELSHOURBAGY, NABIL
;   TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM
;                       RECEPTOR (GPR31A)
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: RATNER & PRESTIA
;     STREET: P.O. Box 980
;     CITY: Valley Forge
;     STATE: PA
;     COUNTRY: USA
;     ZIP: 19482
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/130,749
;     FILING DATE: 07-Aug-1998
;     CLASSIFICATION: UNKNOWN
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: <Unknown>
;     FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;     NAME: PRESTIA, PAUL F
;     REGISTRATION NUMBER: 23,031
;     REFERENCE/DOCKET NUMBER: GP-70513
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 610-407-0700
;     TELEFAX: 610-407-0701
;     TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 960 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-130-749-1

Query Match             13.4%;  Score 139.2; DB 3; Length 960;
Best Local Similarity   51.4%;  Pred. No. 1.6e-30;
Matches  427; Conservative    0; Mismatches 388; Indels 15; Gaps 4;

Qy     62  TGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTCT 121
               ||| || |||  |||| |      | |||||  ||  |||   |||    |
Db     59  TGCTCGGGCCTGGAGTGTGGGCTGGGTCTGGCTCGTGCTGGCGGTGTGGGACCTTCC 118

Qy    122  GCTTCCACATGAAGACCTGAAGCCCAGCACTGTTTACCTTTTCGATTTGGCCGTGGCTG 181
Db    119  TGTTCCGGGTCAGGGGTGTGGAAGCCGTACGGTGTCTCAACCTGTCCCCTGGCTG 178

Qy    182  ATTTCCTCCTTATGATCTGCCGCCTTTTCGGACAGAGACTATTACCTCCAGACGTAGACACT 241
Db    179  ACCTGCTGTTGGCTGCTGCGCGCCTTTCCTGGCCGCCTTCTCACCTGAGCCTCCAGGCTT 238

Qy    242  GGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCCG 301
Db    239  GGCATCTGGGCCGGCCGTGTGGGCTGGGCCCCTGCTTCCTGCTTCCTGCTGCGCCGCAGCG 298

Qy    302  GGAGCATCGTGTTCCTTACGGTGGTGGCTGCCGGACAGGTATTTCAAAGTGGTCCACCCCC 361
Db    299  TGGGGATGGCGCCTTCCTGGCCGGCCGTGTGCTTTGGACCGGTAGTTCACCCCGTGGTTCCACCCTC 358

Qy    362  ACCACGGCCGTGAACACTATCTCCACCCGGGTGGCCGCATCGTGCTGCACCCCCTGTGGG 421
```

```
Db    359  GGCTTAAGGTCAACCTGCTGTCTCCTCAGGCGGCCCTGGGGGTCTCGGGCCTCGTCTGGC 418
Qy    422  CCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAGA 481
Db    419  TCCTGATGGTCGCCCTCCACCTGCCCTGCTCTCATCTCTGAGGCCGCCCAGAACTCCA 478
Qy    482  CGGCCCGTCCTCTGTGGAGAGCTTCATCATGGAGTCGGC------CAATGGCTGCATGACA 535
Db    479  CCAGGTGCCACAGTTTCTACTCCAGGCAGACGGCTCCTTCAGCATCATCTGGCAGGAAG 538
Qy    536  TCATGTTCCAGCTGGAGTCTTTATGCCCCTCGGCATCATCTATTTGCTCCTTCAAGA 595
Db    539  CACTCTCCTGCCTTCAGTTGTTGTCCTCCCCTTGGCCTTCATCCGTGTTCTGCAATGCAGGCA 598
Qy    596  TTGTTTGGAGCCTGAGGCGGAGG---CAGCAGCTGGGCTCGGCCAGATGGCCTCGGATGAAGAAGG 652
Db    599  TCATCAGGGCTCCAGAAAGACTTCCGGAGACCTGAGAAAACAGCCCAAGCTTCAGCGAGCGGG 658
Qy    653  CGACCCGGTTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCCGTGT 712
Db    659  CCCAGGCACTGGTCACCTTGGTGGTGTCGTTTGCTCTGTCCTGCTCCTGTCC 718
Qy    713  CTGCTAGACT-CTATTTCCTGTGGACGGTGCCCTGCGATCCCTCGTCCAT 771
Db    719  TGGCCCAGAGTCCTGATGCACATCTTCCAGAATCTGGGGAGTCTGCAGGGCCCTTTGTGCAG 778
Qy    772  GGGGCCCT-----GCACATAACCCTCAGCTTCACCTACATGAACACAGCATGCTGGATCCCC 826
Db    779  TGGCTCATACCTGGGATGTCACGGCAGCCTCACCTACCTGCACAGTGTGCTCAACCCCG 838
Qy    827  TGGTGTATTATTTTCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTC 876
Db    839  TGGTATACTGCTTCTCCAGCCCCACCTTCAGCGAGCTCCTATCGGAGGGTC 888
```

```
RESULT 11
US-09-130-749-1
; Sequence 1, Application US/09130749
; Patent No. 6031344
; GENERAL INFORMATION:
;   APPLICANT: SHABON, USMAN
;   APPLICANT: ELSHOURBAGY, NABIL
;   TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM
;   TITLE OF INVENTION: RECEPTOR (GPR31A)
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: RATNER & PRESTIA
;     STREET: P.O. Box 980
;     CITY: Valley Forge
;     STATE: PA
;     COUNTRY: USA
;     ZIP: 19482
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/130,749
;     FILING DATE: 07-AUG-1998
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: PRESTIA, PAUL F
;     REGISTRATION NUMBER: 23,031
;     REFERENCE/DOCKET NUMBER: GP-70513
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 610-407-0700
;     TELEFAX: 610-407-0701
;     TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
```

```
;       SEQUENCE CHARACTERISTICS:
;       LENGTH:   960 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;       MOLECULE TYPE:  cDNA
US-09-130-749-1

Query Match         13.4%;      Score 139.2;   DB 3;   Length 960;
Best Local Similarity  51.4%;   Pred. No. 1.6e-30;
Matches  427;   Conservative   0;   Mismatches 388;   Indels 15;   Gaps   4;
```

```
;       GENERAL INFORMATION:
;       APPLICANT:  Behan, Dominic P.
;       APPLICANT:  Chalmers, Derek T.
;       APPLICANT:  Liaw, Chen W.
;       TITLE OF INVENTION: No. 655339-Endogenous, Constitutively Activated Human G Protein
;       TITLE OF INVENTION: Receptors
;       FILE REFERENCE: AREN-0040
;       CURRENT APPLICATION NUMBER: US/09/170,496D
;       CURRENT FILING DATE:  1998-10-13
;       NUMBER OF SEQ ID NOS:  294
;       SOFTWARE: PatentIn version 3.1
;       SEQ ID NO 195
;       LENGTH: 960
;       TYPE: DNA
;       ORGANISM: Homo sapiens
US-09-170-496D-195

Query Match         13.3%;      Score 137.6;   DB 4;   Length 960;
Best Local Similarity  51.3%;   Pred. No. 4.7e-30;
Matches  426;   Conservative   0;   Mismatches 389;   Indels 15;   Gaps   4;
```

RESULT 12
US-09-170-496D-195
; Sequence 195, Application US/09170496D
; Patent No. 6555339

Qy   827 TGGTGTATTATTTTCAAGCCCTCCTTTCCCAAATTCTACAACAAGCTC 876
           ||||||| ||||||||||||||||||| ||||| || ||||||||||||
Db   839 TGGTATACTGCTTCTCCAGCCCCACCTTCAGGAGCTCCTATCGGAGGGTC 888

RESULT 13
US-09-170-496D-59
; Sequence 59, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-59

Query Match   13.1%; Score 136; DB 4; Length 960;
Best Local Similarity   51.2%; Pred. No. 1.4e-29;
Matches   425; Conservative   0; Mismatches 390; Indels 15; Gaps 4;

Qy   62 TGCTCATTGTGGCCTTGTGCTGGCCGCACTAGGCAATGGGTCGCCCTGTGTGGTTCT 121
           ||||||||||||| ||||||| ||||||||||||
Db   59 TGCTGGGGCTGGAGTGTGGGCTGGCTCGGGCAACGCGGTGGCGGTGGACCTTCC 118

Qy   122 GCTTCCACATGAAGACCTGGAAGCCCAGCACAGTGTTTCACCTTTCAATTCGGCCGTGGGCTG 181
Db   119 TGTTCCGGGTCAGGGTGTGAAGCCGTACGCGTGTCTTCACCTGGCTGGCCCTGGCTG 178

Qy   182 ATTTCCTCCTTATGATCTGGCCTGCCTTTCCGACAGACTATTACCTCAGACGTAGACACT 241
Db   179 ACCTGCTGTTGGCTGCGGCCTGCCTTCTCCTGCCGTCTTACCTGAGCCTCCAGGCTT 238

Qy   242 GGGCTTTTGGGGACAATTCCCTGCCCAGTGGGGCTCTCACGTTGGCCATGAACAGGGCCG 301
Db   239 GGCATCCTGGGGCCGTGTGGCCTGTGCTGGGCCGTGCTGCTCTGCTGCCTCAGCCGCAGCG 298

Qy   302 GGAGCCATGCGTGTTCCTTACCGGTGGTGGCTGGCACAGGTATTTCAAAGTGGTCACCCCC 361
Db   299 TGGGGATGGCCGGCCTTCCTGGCCGCGTGGCCTTTGGACCGGTACCTCCGTGTGGTTCCACCCTC 358

Qy   362 ACCACGCGGTGAACACTATCTCCACCCCGGTGGCGGCTGGCACCCTGTGGG 421
Db   359 GGCTTAAGGTCAACCTGGTCTCCTGCTGCCCCTGGGGGGTCTGGGGCCTTGGTCTGGC 418

Qy   422 CCCTCGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCGTGCAAGAGA 481
Db   419 TCCTGATGGTCTCCGCCTCCACCTGCTGCTCATCGCGCCCGGCCGCCCAGAAGACTCCA 478

Qy   482 CGGCCGTCATCCTGGAGAGCTTCATCATGAGTCGGC------CAATGGCTGGCATGACA 535
Db   479 CCAGGTGCCACAGTTTCTACTCCAGGCAGACGCGGTCCTCTTCAGCATCATCTGGCAGGAAG 538

Qy   536 TCATGTTCCAGCTGAGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGTCTCCTTCAAGA 595
Db   539 CACTCTCCTGCCTGCCTTCAGTTGTTCTTGCCTCATCGTGTTCTGCAATGCAGGCA 598

Qy   596 TTGTTTGGAGCCTGAGGCGGAGG---CAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGG 652
Db   599 TCATCAGGGGCTTCTTCCAGAAAAAGACTCCCGGAGCTCCCAAGCTTCAGCGGG 658

Qy   653 CGACCCGGGTTCATCATGGTGGTGGCAATTGTGTTCATCATCCTGCCCAGCCGTGT 712
Db   659 CCCAGCCACCTGTCAACCTTGGGTGTCGTTTGCTCGTGCTTTCTCTGCCCTGCTCC 718

Qy   713 CTGCTAGACT-CTATTTCCTCTGGACGGTGGCCCTCGAGTGCTGCGATCCCTGTGTCCAT 771
Db   719 TGGCCCAGAGTCCTGATGCACATCTTCCAGAATCTCCAGCAGGGCTGCAGGGCCCTTGTGCAG 778

Qy   772 GGGGCCCT-----GCACATAACCCTCAGCTTCACCTACATAGAACAGCATGCTGGATCCCC 826
Db   779 TGGCTCATACCTCGGATGTCACGGGCAGCCTCACCTACTGCACAGTGTCGTCAACCCCG 838

Qy   827 TGGTGTGTATTATTTTCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTC 876
Db   839 TGGTATACTGCTTCTCCAGCCCCACCTTCAGGAGCTCCTATCGGAGGGTC 888

RESULT 14
US-09-170-496D-31
; Sequence 31, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protei
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-31

Query Match   10.5%; Score 109; DB 4; Length 1020;
Best Local Similarity   47.1%; Pred. No. 9.6e-22;
Matches   377; Conservative   0; Mismatches 415; Indels 9; Gaps 1;

Qy   48 GGTGATGCGCCGCTGCTCATTGTGGCCTTGTGCTGGCCGCACTAGGCAATGGGGTCGC 107
Db   96 GCTGTTCGCCTCCTTCTACTTCTCTGGATTTTATCCTGGCTTTAGTTGGTGGCAATACCCTGGC 155

Qy   108 CCTGTGTGGTTCTGCTTCCACATGAAGACCTGGAAGCCCAGCACAGTGTTTACCTTTCAA 167
Db   156 TCTGTTGGCTTTTCATCCGAGACCACAAGTCCGGGACCCCCAACGTGTCCTGATGCA 215

Qy   168 TTTGGCGTGGCTGGCTGATTTCCTCCTATGATCTGGCGACAGACTATTACCT 227
Db   216 TCTGGCCGTGGCGACTTGTGCGTGCTGGTCTGCTTGTCGTCTTACCACTT 275

Qy   228 CAGACGTAGACACTGGGCGTTTGGGACATTCCCTGCCCAGTGGGGCTCTCACGTTGGC 287
Db   276 CTCTGGGAACCACTGGCCCATTGGGAAATCGCATCGCCATGCCCGACCGTTCCTGGCC 335

Qy   288 CATGAACAGGGCCGGGAGCATGTGTTCCTTACCGGTGGTGGCTCGCGGGCCCGGGTGGGCG 347
Db   336 CCTCAACATGTACGCCAGCATCCAGCCATCCTCCTCCTCACCTGACCTGCACCGGTTCT 395

Qy   348 AGTGGTCCACCCCACCAGCGGTGAACACTATCTCCACCGTCGGCGGCCGTCAGGTATTTCAA 407
Db   396 CATTGTGCACCCGGTCAGTCCCTCAGCTCCTGGAACAGTGTATCTTTTGCTGGAACCATCT 455

Qy   408 CTTGCACCCTGTGGGCCCTGGTTCTGCTTCATCCTCGGATGGCCAATGCAGGCA 467
Db   456 TGCCTTCCTCCTGGTGGTGGTGTGCCCCGCTGCTGGTGAGCCCACAGAC 515

Qy   468 CTTGCGTGCAAGAGACGGCCGTCTTCATCATCGAGAGCTTCATCATGAGTCGGCCAATGGCTG 527
Db   516 CGTGCAGGACCAACCACACGGTGGTCTCGCCTCGCAGCTGTACCGGGAGAAGGCCTCCACCA 575

Qy   528 GCAATGACATCATGTTCCAGCTGAGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTC 587

Matches 377; Conservative 0; Mismatches 415; Indels 9; Gaps 1;

```
Qy    48 GGTGATGCCGCCGCTGCTCATTGTGGCCTTTGTCGTCCGGCCGCACTAGGCAATGGGTCGC 107
Db   795 GCTGTTGCCTCCTTCACCTTCTGGATTTTATCCTGGCTTTTAGTTGGCAATACCCTGGC   854

Qy   108 CCTGTGTGTTTCGTTCCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAA 167
Db   855 TCTGTGGCTTTTCATCCGAGACCACCAAGTCCGGGACCCCGGCCAACGTGTTCCTGATGCA 914

Qy   168 TTTGGCCGTGGCTGATTTCCTCTTATGATCCTCCCTGCCTTTCGGACAGACTATTACCT   227
Db   915 TCTGGCCGTGGCCCGACTTGTCGTGCGTGGTCGTCGTCCGCCTGGTCTACCACTT       974

Qy   228 CAGACGTAGACACTGGGCTTTTGGGGACATTCCCTGCCAGTGGGGCTCTTCACGTTGGC   287
Db   975 CTCTGGGAACCACTGGCCATTTGGGAAAATGCATGCCGTTCCACCGGCTTCCTCTTCTA   1034

Qy   288 CATGAAACAGGGCCGGGAGCATCGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAA   347
Db  1035 CCTCAACATGTACGCCAGCATCTACTTCTCCACCTGACCTGACCGTTCCTGGC        1094

Qy   348 AGTGGTCACCCCCACCACGCGGTGAAACATATCTCCACCCGGGTGGCGGCTGGCATCGT   407
Db  1095 CATTGTGCACCCGGTCAAGTCCCCAAGCTCCGCAGGCCCCTCTACGCACACTGGCCTG    1154

Qy   408 CTGCACCCTGTGGGCCTTGGTCATCCTGGGAAACAGTGTATCTTTTGCTGGAGAACCATCT 467
Db  1155 TGCCTTCCTGTGGGTGGTGGCCTGGCCATGGCCCGCTGGTGGTGAGCCCACAGAC       1214

Qy   468 CTGGCTGCAAGAGACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTG  527
Db  1215 CGTGTCAGACAGCAACCAACACACGGTGGTCTGCCTGCAGCTGTACCGGGAGAAGGCCTCCCACCA 1274

Qy   528 GCATGACATCATGTTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTTATTTTGCTC   587
Db  1275 TGCCCTGGTGTCCCTGGCCAGTGGCCTTCACCTTCCGTTCATCACCACGGTCACCTGCTA  1334

Qy   588 CTTCAAGATTGTTTGGAGCCTGGAGGCGGAGGCAGCAGGCTGGCCAGACAGCCTCGGATGAA 647
Db  1335 CCTGCTGATCATCCGCAGCCTGCGGCAGCCTGCGTGTGGAGAGAGCGCCTCAAGACCAA    1394

Qy   648 GAAGGCGACCCGGTTCATCATGGTGGTGGCAATTGT-------GTTCATCACATGCTA    698
Db  1395 GGCAGTGCGCATGATGCGCCATAGTGCTGGCCATCTTCCTGGTTCGTGCCCTACCA      1454

Qy   699 CCTGCCCAGCGTGTCTGTCTAGACTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGA 758
Db  1455 CGTCAACCGGTCCGGTCAGTGGCTTCACTTACCGCAGCCATGGGCCTTCCTCCGCGCCACCCCA 1514

Qy   759 TCCCTCTGTCCATGGGGCCCTGCACATAACCCTCAGCTTCACTACATGACAGCATGCT    818
Db  1515 GCGCATCCTGGCCCTGCCAAACCGGCATCACCTCCTCGCCTCACCAGCCTCAACGGGGGCACT 1574

Qy   819 GGATCCCCTGGTGTATTATTT 839
Db  1575 CGACCCCATCATGTATTTCTT 1595
```

```
Db   576 TGCCCTGGTGTGCCCTCGGCAGTGGCCTTCACCTTCCCGTCATCACCACGGTCACCTGCTA 635
Qy   588 CTTCAAGATTGTTTGGAGCCTGGAGGCGGAGGCAGCAGGCTGGCCAGACAGCTCGGATGAA  647
Db   636 CCTGCTGATCATCCGCAGCCTGCGCGGCCTGCGTGTGGAGAGAGCGCCTCAAGACCAA     695
Qy   648 GAAGGCGACCCGGTTCATCATGGTGGGCAATTGT-------GTTCATCACATGCTA      698
Db   696 GGCAGTGCGCATGATGCGCCATAGTGCTGGCCATCTTCCTGGTCGTTGCGTCCCTACCA    755
Qy   699 CCTGCCCAGCGTGTCTGTCTAGACTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGA 758
Db   756 CGTCAACCGGTCCGGTCAGTGGCTTCACTACCGCAGCCATGGGCCTTCCTCCGCGCCACCCA 815
Qy   759 TCCCTCTGTCCATGGGGCCCTGCACATAACCCTCAGCTTCACTACATGACAGCATGCT    818
Db   816 GCGCATCCTGGCCCTGCCAAACCGGCATCACCTCCTCGCCTCACCAGCCTCAACGGGGGCACT 875
Qy   819 GGATCCCCTGGTGTATTATTT 839
Db   876 CGACCCCATCATGTATTTCTT 896
```

RESULT 15
US-09-016-434-1484
; Sequence 1484, Application US/09016434
; Patent No. 6500938
GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; LIBRARY: GENBANK
; CLONE: 9992699
US-09-016-434-1484
; INFORMATION FOR SEQ ID NO: 1484:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1900 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9992699
US-09-016-434-1484

Query Match   10.5%; Score 109; DB 4; Length 1900;
Best Local Similarity 47.1%; Pred. No. 1.3e-21;

Search completed: July 3, 2004, 08:44:43
Job time : 110 secs

```
OM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:         July  3, 2004, 06:00:01 ; Search time 3245 Seconds
                                           (without alignments)
                                           9552.217 Million cell updates/sec

Title:          US-10-076-260-1
Perfect score:  1038
Sequence:       1 atgtacaacgggtcgtgctg.........cccacattgttgagtggcac 1038
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters:       55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :    EST:*
          1:  em_estba:*
          2:  em_esthum:*
          3:  em_estin:*
          4:  em_estmu:*
          5:  em_estov:*
          6:  em_estpl:*
          7:  em_estro:*
          8:  em_htc:*
          9:  gb_est1:*
         10:  gb_est2:*
         11:  gb_htc:*
         12:  gb_est3:*
         13:  gb_est4:*
         14:  gb_est5:*
         15:  em_estfun:*
         16:  em_estom:*
         17:  em_gss_hum:*
         18:  em_gss_inv:*
         19:  em_gss_pln:*
         20:  em_gss_vrt:*
         21:  em_gss_fun:*
         22:  em_gss_mam:*
         23:  em_gss_mus:*
         24:  em_gss_pro:*
         25:  em_gss_rod:*
         26:  em_gss_phg:*
         27:  em_gss_vrl:*
         28:  gb_gss1:*
         29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1038	100.0	1372	11	AF345568	AF345568 Homo sapi
	2	722.8	69.6	3028	11	AK029064	AK029064 Mus muscu
	3	711.8	68.6	3283	11	AK046700	AK046700 Mus muscu
c	4	705.4	68.0	792	14	CD559495	CD559495 AGENCOURT
c	5	676	65.1	786	14	CD559498	CD559498 AGENCOURT
c	6	672.8	64.8	793	14	CD559497	CD559497 AGENCOURT
c	7	669.4	64.5	790	14	CD559496	CD559496 AGENCOURT
	8	619.8	59.7	773	14	CD559653	CD559653 AGENCOURT
	9	600	57.8	748	14	CD559650	CD559650 AGENCOURT
	10	591.4	57.0	743	14	CD559651	CD559651 AGENCOURT
	11	584.4	56.3	731	14	CD559652	CD559652 AGENCOURT
	12	558.6	53.8	1004	14	CD246157	CD246157 AGENCOURT
c	13	453.8	43.7	594	29	CE432388	CE432388 tigr-gss-
	14	443.8	42.8	635	14	CB576966	CB576966 AMGNNUC:U
	15	370.8	35.7	2146	11	BC027965	BC027965 Homo sapi
	16	367.6	35.6	2068	11	BC056419	BC056419 Homo sapi
	17	367.6	35.4	2068	11	BC047891	BC047891 Homo sapi
	18	324.6	31.3	1201	9	AL545172	AL545172 AL545172
	19	317.2	30.6	960	9	AL546894	AL546894 AL546894
	20	309.2	29.8	1057	12	BM918711	BM918711 AGENCOURT
	21	309	29.8	801	12	BI837965	BI837965 603083801
	22	307	29.6	1076	12	BM923028	BM923028 AGENCOURT
	23	301.4	29.0	1063	12	BM920158	BM920158 AGENCOURT
	24	300.8	29.0	472	14	CB730514	CB730514 AMGNNUC:N
	25	289.4	27.9	1201	9	AL554198	AL554198 AL554198
	26	278.6	26.8	876	12	BI820995	BI820995 603035649
	27	270.4	26.1	734	12	BI908162	BI908162 603067472
	28	267.2	25.7	752	12	BMC08116	BMC08116 603618002
	29	255.8	24.6	367	10	BB873518	BB873518 BB873518
	30	254.6	24.5	660	10	BB614132	BB614132 BB614132
	31	253.2	24.4	367	10	BB873597	BB873597 BB873597
	32	251	24.2	372	13	BY234567	BY234567 BY234567
	33	236.6	22.8	686	13	BY747695	BY747695 BY747695
	34	230.4	22.2	921	13	BQ706073	BQ706073 BB849390
	35	230	22.2	422	10	BB849390	BB849390 BB849390
	36	224.2	21.6	690	14	CD690497	CD690497 EST7020 h
	37	220.2	21.2	645	14	CD469236	CD469236 LeukoS2_2
	38	218.2	21.0	550	10	BF074336	BF074336 221723 MA
	39	217.2	20.9	1082	13	BQ710754	BQ710754 AGENCOURT
	40	217	20.9	922	13	BU146785	BU146785 AGENCOURT
	41	216.8	20.9	612	14	CD536045	CD536045 LeukoN6_2
	42	215.8	20.8	636	12	BI489649	BI489649 603032118
	43	215.2	20.7	614	14	CD468367	CD468367 LeukoS3_2
	44	213.8	20.6	514	13	BX279654	BX279654 BX279654
	45	213.8	20.6	532	10	BF230383	BF230383 251977 BA

ALIGNMENTS

```
RESULT 1
AF345568
LOCUS       AF345568                1372 bp    mRNA    linear    HTC 02-MAY-2003
DEFINITION  Homo sapiens putative chemokine receptor (FKSG80) mRNA, complete
            cds.
ACCESSION   AF345568
VERSION     AF345568.1  GI:13517963
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1372)
  AUTHORS   Alberta,E.O., Petersen,R.H., Hughes,K.W. and Lechner,B.
  TITLE     Miscellaneous notes on Pleurotus
  JOURNAL   Persoonia 18, 55-69 (2002)
REFERENCE   2  (bases 1 to 1372)
  AUTHORS   Wang,Y.-g. and Gong,L.
  TITLE     Molecular cloning of FKSG80, a novel gene encoding a putative
            chemokine receptor
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1372)
  AUTHORS   Wang,Y.-g.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2001) Beijing FENGKESHENG Function Gene
            Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District,
            Beijing 100050, P.R. China
```

```
FEATURES             Location/Qualifiers
source               1..1372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
gene                 1..1372
                     /gene="FKSG80"
CDS                  142..1182
                     /gene="FKSG80"
                     /codon_start=1
                     /product="putative chemokine receptor"
                     /protein_id="AAK29071.1"
                     /db_xref="GI:13517964"
                     /translation="MYNGSCCRIEGDTISQVMPPLLIVAPVLGALGNGVALCGFCFHM
                     KTWKPSTVYLFNLAVADPLLMICLPFRTDYILRRRHWAPGDIPCRVGLFTLAMNRAGS
                     IVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCTLMALVIIGTVYLLLENHLCVQE
                     TAVSCESFIMESANGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKA
                     TRPIMVVAIVFITCYLPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPL
                     VYYFSSSFPKFYNKLKICSLKPKQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQS
                     DGQWDPHIVEWH"
ORIGIN
```

```
Query Match        100.0%;  Score 1038;  DB 11;  Length 1372;
Best Local Similarity 100.0%;  Pred. No. 2.6e-268;
Matches 1038;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 AIGTACAACGGGTCGTGCTGCGCCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 ATGTACAACGGGTCGTGCTGCGCCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG  201

Qy    61 CTGCTCAATTGTGGCCTTTGTCTGGGCGCACTAGGCAATGGGCGTGCCCTGTGTGGTTTC  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 CTGCTCAATTGTGTGGCCTTTGTCTGGGCGCACTAGGCAATGGGCGTGCCCTGTGTGGTTTC  261

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  321

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGCATTACCTCAGACTAGACAC  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGCATTACCTCAGACTAGACAC  381

Qy   241 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  441

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   442 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC  501

Qy   361 CACCACGCGGTGAACACTATCTCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 CACCACGCGGTGAACACTATCTCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG  561

Qy   421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCTGCAAGAG  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCTGCAAGAG  621

Qy   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG  681

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTGGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 TTCCAGCTGGAGTTCTTTATGCCCCTGGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  741

Qy   601 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG  801

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 TTCATCATGGTGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGGTGTCTGCTAGA  861

Qy   721 CTCTATTTCCTCTGGACGGTGCCTGAGTGGCCTGGATCCCTGTCCATGGGGCCCTG  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 CTCTATTTCCTCTGGACGGTGCCTGAGTGGCCTGGATCCCTGTCCATGGGGCCCTG  921

Qy   781 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGTGTGTATTATTT  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTGTGTGTATTATTT  981

Qy   841 TCAAGCCCTCCTTTCCCAAATTCTACAACAAGCTCAGTCGAGTCTGAAACCCAAG  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 TCAAGCCCTCCTTTCCCAAATTCTACAACAAGCTCAGTCGAGTCTGAAACCCAAG 1041

Qy   901 CAGCCAGGAGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1042 CAGCCAGGAGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 1101

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1102 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1161

Qy  1021 CACATTGTTGAGTGGCAC 1038
           ||||||||||||||||||
Db  1162 CACATTGTTGAGTGGCAC 1179
```

```
RESULT 2
AK029064
LOCUS       AK029064    3028 bp    mRNA    linear    HTC 18-SEP-2003
DEFINITION  Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
            library, clone:4732488A18 product:similar to PUTATIVE CHEMOKINE
            RECEPTOR (G PROTEIN-COUPLED RECEPTOR) [Homo sapiens], full insert
            sequence.
ACCESSION   AK029064
VERSION     AK029064.1  GI:26325037
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,B., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system-384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
```

REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 3028)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..3028
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM DB:4732488A18"
                     /db_xref="MGI:2391098"
                     /db_xref="taxon:10090"
                     /clone="4732488A18"
                     /tissue_type="skin"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="10 days neonate"
     CDS   404..1435
                     /note="unnamed protein product; putative
                     similar to PUTATIVE CHEMOKINE RECEPTOR (G PROTEIN-COUPLED
                     RECEPTOR) [Homo sapiens] (SPTR|Q9BXC0, evidence: FASTY,
                     80.9%ID, 100%length, match=1039)"
                     /codon_start=1
                     /protein_id="BAC26273.1"
                     /db_xref="GI:2632503B"
                     /translation="MDNGSCCLIEGEPISQVMPPLLILVFVLGALGNGIALCGFCFHM
                     KTWKSSTYLFNLAVADFLMICLPLRTDYYLRRRHWIFGDIACRLVLFKLAMNRAGS
                     IVFLTVVAVDRYFKVVHPHHMVNAISNRTAAATACVLWTLVILGTVYLLMESHLCVQG
                     TLSSCESFIMESANGWHDVMFQLEFFLPLIIILFCSVNVVWSLRRRQQLTRQARMRRA
                     TRFIMVVASVFITCYLPSVLARLYFLWTVPTSACDPSVHTALHVTLSFTYINSMLDPL
                     VYYFSSPSLPKFYTKLTICSLKPKRPGRTKIRRSEEMPISNLCSKSSIDGANRSQRPS
                     DGQWDLQVC"
     polyA_signal   3011..3016
                     /note="putative"
     polyA_site   3028
                     /note="putative"
ORIGIN

```
Db    404 ATGGACAACGGGTCGTGCTGTCTCATCGAGGGGGAGCCCATCTCCCAAGTGATGCCTCCT 463
          |||||||||| ||||||||||||||||||||||||||||||||||||||  ||||||||||
Qy     61 CTGGTCATTGTGGCCTTGTGGCGGGCCACTAGGCAATGGGGTCGCCCTGTCGTTGGTTTC 120
          |||||| ||  ||||||||||||||||||||||||||  ||  |||||||||||||||||
Db    464 CTACTCATCCTGGTCTTCGTGCTTGGCGCCCTTGGGCAACGGCATAGCCCTGTGCGGCTTC 523

Qy    121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    524 TGCTTCCACATGAAGACCTGGAAGTCAAGCACTATTACCTTTTCAACTGGCTGTGGCC 583

Qy    181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGAGTAGACAC 240
          || |||||||| || |||||||||||||||||||||||||||||||||||||||||||||
Db    584 GATTTTCCTCCTCATGATCTGCTTACCCCTTCGGACACAGACTACTACCTCAGAGACAGAC 643

Qy    241 TGGGGCTTTGGGGACATTCCCTGCCAGATGGGGTCTCTTCACGTTGGCCATGAACAGGGCC 300
          |||||| ||||||||| || ||||||||||||||||||| |||||||||||||||||||||
Db    644 TGGGATTTTGGAGATATCGCCTGCCGTGTCGGTTCCTCTTCAGCTGCCATGAATAGGGCC 703

Qy    301 GGGAGCAATGCTGTGTTCCTTACGGTGGTGGCCGGGACAGTATTTCAAAGTGGTCCACCCC 360
          |||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    704 GGGAGCAATTGTGTCTCCTCACTGTGGCGGTGGCTGGATAGGTATTTCAAAGTGGTCCACCCC 763

Qy    361 CACCACGCGGTGAACACTATCTCCACCGGCGGTGGCCGGGATCGTCTGCACCCTGTGG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 CACCATATGTGAATGCCATCTCCAACCGGACTGCCCGCGGCCTGTGTGTCCTCTGG 823

Qy    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGTGGAGAACCATCTCTGCGGTGCAAGAG 480
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    824 ACTTTGGTCATCTTGGGACTGTGTATCTTCTCTGATGGAGAGTCACCTGTGTGCAGGGG 883

Qy    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    884 ACACTGTCGTCCTGTGAGAGCTTCATCATGGAGTCAGCCAACGGGTGGCACGATGTCATG 943

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTGTCGGTCCTTCAAGATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    944 TTCCAGCTGGAGTTCTTCCTGCCCCTGACAATCATCTTGTTCTGTTCGGTCAACGTTGTT 1003

Qy    601 TGGAGCCTGAGGCGGAGGCAGCTGACCAGGCAGGCTCGGATGAGAGAGGCGACCCGG 660
          ||||||||||||||||||||||||||||||||||  ||||||||||| ||||||||||
Db   1004 TGGAGCCTGAGACGGAGGCAGCTGACCAGACAGGCTCGGATGAGGAGGGCCACCCGG 1063

Qy    661 TTCATCATGGTGGTGGCAATTGTGTCATCACATGCCAACATGCCCAGCGTGTCTGCAGA 720
          |||||||||||||||||||||||||||||||  ||||||||||||||||||||| |||||
Db   1064 TTCATCATGGTGGTCGCTTCGTGTGTTCATCACGTGTTACCTGCCCAGCGTGCTGGCTAGG 1123

Qy    721 CTCTCATTCCTCGGACGGTGCCCTGGACGGTGCCCGAGTCGTGCATGGGGGCCCTG 780
          |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db   1124 CTCTACTTCCTCTGGACGGTGCCCACTAGTGCTTGCGACCCCTCGTGTCACACAGCCCTC 1183

Qy    781 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTCGATCCCCTGGTGTATTATTT 840
          |||| ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   1184 CACGTCACCCTGAGCTTCACCTACATCAACAGCATGCTCGATCCCCTGGTGTATTATTACTTC 1243

Qy    841 TCAAGCCCCTCCTTTCCAAATTCCAGTGCTGAACCCAAG 900
          |||||||||| |  ||  ||||  ||||||||||||||||
Db   1244 TCAAGCCCCTCGCTCCCCAAATTCTACAACAAGCTCAAATCTGAAGCCAAA 1303

Qy    901 CAGCCAGGACACTCAAATGCCAATTCGAACCTCGTCGC 960
          |||||||||||||||||| ||||||||||||||||||||
Db   1304 CGCCCAGGACGACGCGCGGAGGTCAGAAGAGATGCCAATTTCGAACCTCTGCAGT 1363

Qy    961 AGGAGTTGCATCAGTGTGCAAATAGTTTCAAAGCCAATGGGATCCC 1020
          |||| ||| |||||||||||||| ||||||||||||||||||||||
Db   1364 AAGAGCTCCATCGATGGGCACAAATCGTCCCAGAGGGCAGTGGGATCTC 1423

Qy   1021 CA 1022
          ||
Db   1424 CA 1425
```

RESULT 3

```
Query Match       69.6%;   Score 722.8;  DB 11;  Length 3028;
Best Local Similarity   81.7%;   Pred. No. 4e-183;
Matches   835;   Conservative   0;  Mismatches  187;  Indels   0;  Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
```

```
AK046700
LOCUS       AK046700            3283 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length
            enriched library,clone:B430318H17 product:similar to PUTATIVE
            CHEMOKINE RECEPTOR (G PROTEIN-COUPLED RECEPTOR) [Homo sapiens],
            full insert sequence.
ACCESSION   AK046700
VERSION     AK046700.1  GI:26091668
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3283)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,K., Sakai,C., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
```

```
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..3283
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:B430318H17"
                     /db_xref="MGI:2412217"
                     /db_xref="taxon:10090"
                     /clone="B430318H17"
                     /sex="male"
                     /tissue_type="adipose"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="4 days neonate"
     misc_feature    420..1462
                     /note="putative
                     similar to PUTATIVE CHEMOKINE RECEPTOR (G PROTEIN-COUPLED
                     RECEPTOR) [Homo sapiens] (SPTR|Q9BXC0, evidence: FASTY,
                     80.9%ID, 100%length, match=1039)"
     polyA_signal    3262..3267
                     /note="putative"
     polyA_site      3283
                     /note="putative"
ORIGIN
Query Match              68.6%;    Score 711.8;   DB 11;   Length 3283;
Best Local Similarity    81.6%;    Pred. No. 3.8e-180;
Matches   835;   Conservative   0;  Mismatches 187;  Indels   1;  Gaps   1;

Qy      1 ATGTACAACGGGTCGTGCTGCCCGCATGAGGGGGACACCATCTCCAGGTGATGCGCCG  60
          ||| |||||||||||||| ||||||| ||||||| |||| ||||||||||||| ||||||
Db    420 ATGGACAACGGGTCGTGCTGCTCATGCGGGGAGGGGAGCCCATCTCCAAGTGATGCCTCCT 479

Qy     61 CTGCTCATTGTGGCCTTTGTGTCTGGGCGCACTAGGCAGGGTCGCCCTGTGTGGTTTC 120
          |||||||| |||||| |||||| ||||| | || ||||||||  ||||| ||||||||
Db    480 CTACTCATCCTGGCGTTGGCCGTCTGGCCCGGCAACGGCATAGCCCTGTGTGCGGGCTTC 539

Qy    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTACCTTTTCAATTGGCGTGGCT 180
          |||||| |||||||| ||||| ||| ||||| |||| || |||| ||||| ||||| |||
Db    540 TGCTTTCACATGAAGACCTGGAAGTCAAGCACTATTTACCTTTTCAACTTGGCGTGGCC 599

Qy    181 GATTCCTCCTATGATCTGCCTGCTTTTCGGACAGACTATTACCTCCAGACGTAGACAC 240
          ||||| ||| ||||| |||| |  ||| |||| ||||  ||| |||||||||| |||||
Db    600 GATTTTTGGAGATATCGCCTGGTCCTCCTCTTCAAGCTGTATTTCAACTTGGCCATGAATAGGGCC 659

Qy    241 TGGGCTTTGGGGACATTCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGACAGAGGCC 300
          |||||| ||||| |||| |||  ||||||||||||| |||||||||||| |||| |||||
Db    660 TGGATTTTGGAGATATCCTGGTCCTCCTCTTCAAGCTGTATTCAAGCTGGGCCATGAATAGGGCC 719

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCTGCGGAGTTATTCAAAGTGGTCCACCCC 360
          ||||| ||||||  ||||| |||| || ||| |||||||| |||| ||||||  |||||
Db    720 GGGAGCATTGTCTTCCTCCACTGTGGCTGTGGGCTGTGGATAGGTATTCAAAGTGGCCACCCC 779

Qy    361 CACCACGCCGGTGAACACTATCTCCACCCGGGGTGGGCGGCTGCATCGTCTGCAGGAGGCC 420
          ||||| |||| |||||||||||| ||||| |||||| ||  ||| |||| ||||| |||||
Db    780 CACCATATGGTGAATGCCATCTCCAAGCCGGAATGCGGGACTGCGCCCGCCGTGTCCTCTGG 839

Qy    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
          ||||| ||||||||||||||||||| |||||||||  ||||||| ||| |||| |||||||
Db    840 ACTTTGGTCATCTTGGGAGACTGTGTATCTTCTGATGGAGAGTCACCTGTGTGTGCAGGGG 899

Qy    481 ACGGCCGTCTCCGTGGAGAGCTTCATCATGACGTCGGCAACATGGCTGGCATGACATCATG 540
          |||| || |||||| || ||||| |||||| |   | ||| |||| ||||| |||||||||
Db    900 ACACTGCGTCCGTGGAGAGCTTCATCATGATGGAGAGTCAGCCAACGCGGGTGGCACCATGTCATG 959

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTGCGCATCACATCTATTTGCTCCTTCAAGATTGTT 600
          |||||||||| ||||| |||  |||||||||| |||| ||||||| || |||||||||||
```

Db   960 TTCCAGCTGGAGTTCTTCCGCCCCTGACAATCATCTTGTTCTGCTCGGTCAACGTTGTT 1019

Qy   601 TGGAGCCTGAGCGGAGGCAGCAGCAGCTGGCCAGACAGGCTCGGATGAAGGCGA-CCCG 659
                ||||||||||| ||||||||| |||||| |||||| |||||||||||||||| ||||
Db   1020 TGGAGCCTGAGCACGAGGAGGCAGCAGCTGACCAGACAGGCTCGGATGAAGGGGCCACCCCG 1079

Qy   660 GTTCATCATGGTGGCGGCCAATGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAG 719
                ||||||||||||||||| ||||||||||||| ||||||||||||||||||||||||||
Db   1080 GTTCATCATGGTGGCGTGGCTTCGTGTTCATCACGTGTTCATCACGTGCTGGCTAG 1139

Qy   720 ACTCTATTTCCTTCGGACGGGTGCCCTCGAGTGCCTCGGATCCCTCTCGTCCGTGGGGCCCT 779
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 GCTCTACTTCCTTCGGACGGGTGCCCACTAGTGCCTGTGACCCCCTCGTCCACACAGCCCT 1199

Qy   780 GCACATAACCCTCAGCTTCACCTACTACATGAACAGCATGCTGGATCCCCTGGTGTATTATTT 839
                | ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1200 CCACGTCACCCTGAGCTTCACCTACCTGAACAGCATGCTGGATCCCCTTGTATATTACTT 1259

Qy   840 TTCAAGCCCCCTTCCTTCCCAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAA 899
                ||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db   1260 CTCAAGCCCCCTCGCTCCCCCAAATTCTCACACCAAGCTCACAATCTGCAGCCTGAAGCCCAA 1319

Qy   900 GCAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCG 959
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 ACGCCCAGGACGCACGAAGACGCGGAGGTCAGAGAGAGATGCCAATTTCGAACCTCTGCAG 1379

Qy   960 CAGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCC 1019
                |||||||||||||||||||| |||| |||||||||||||||||||||||||||||||||
Db   1380 TAAAGAGCTTCATCGATGGGCAAATCGTTCCCAGAGCCCATCTGACGGGCAGTGGGATCT 1439

Qy   1020 CCA 1022
                |||
Db   1440 CCA 1442

RESULT 4

CD559495/c

LOCUS   CD559495   792 bp   mRNA   linear   EST 19-NOV-2003
DEFINITION AGENCOURT_14497073 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:6971820 5', mRNA sequence.
ACCESSION   CD559495
VERSION   CD559495.2 GI:38453559
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 792)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   On Jun 10, 2003 this sequence version replaced gi:31585563.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
            cDNA Library Preparation: Bhat Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK2 row: C column: 11
            High quality sequence start: 9
            High quality sequence stop: 726.
FEATURES   Location/Qualifiers
     source   1..792
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6971820"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presv.dat
                     a Note: this is a NIH_MGC library."

ORIGIN

Query Match   68.0%;   Score 705.4;   DB 14;   Length 792;
Best Local Similarity 98.8%;   Pred. No. 9.6e-179;
Matches 718;   Conservative 0;   Mismatches 8;   Indels 1;   Gaps 1;

Qy   312 GTTCCTTACGGTGGTGCGGCAGCAGTATTTCAAAGTGGTCACCCCACCACGCGGGT 371
                || ||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db   758 GTGTCCTTACGGTGGTGCGGCAGCACAG-ATTCNAAAGGGTCACCCCACCACGCGGGT 700

Qy   372 GAACACTATCTCCACCCGGGTGCGGCTGGGCATGCTGCTCGTCTCGGTGCTCGGTGGTCAT 431
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   699 GAACACTATCTCCACCCGGGTGCGGCTGGGCATCGTCTGCTCGGTGCTCGGTGGTCAT 640

Qy   432 CCTGGAACAGTGTATCTTTGCTGAGAACCATCTCGGTGCAAGAGACGGGCCGTCTC 491
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   639 CCTGGAACAGTGTATCTTTGCTGAGAACCATCTCGGTGCAAGAGACGGGCCGTCTC 580

Qy   492 CTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATGTTCCAGCTGGA 551
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   579 CTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATGTTCCAGCTGGA 520

Qy   552 GTTCTTTATGCCCCTGCGCATCATCATCTTATTTGCTCCTTCAAGATTGTTTGGAGCCTGAG 611
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   519 GTTCTTTATGCCCCTGCGCATCATCATCTTATTTGCTCCTTCAAGATTGTTTGGAGCCTGAG 460

Qy   612 GCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGGCACCCGGTTCATCATGT 671
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   459 GCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGGGCCCGGTTCATCATGT 400

Qy   672 GGTGGCAATTGTGTTCATCACGTCACCTGCCCAGGTGTCTGCTAGACTCTATTTCCT 731
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   399 GGTGGCAATTGTGTTCATCACGTCACCTGCCCAGGTGTCTGCTAGACTCTATTTCCT 340

Qy   732 CTGGACGGGCCCTCGAGTGCCTCGTCCTGTCCATGGGGCCCTGCACATAACCCT 791
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   339 CTGGACGGGCCCTCGAGTGCCTCGTCCTGTCCATGGGGCCCTGCACATAACCCT 280

Qy   792 CAGCTTCACCTACATGAACAGCATGCTGGATCCCGTGTATTATTTTTCAAGCCCCTC 851
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   279 CAGCTTCACCTACATGAACAGCATGCTGGATCCCGTGGTGTATTATTTTTCAAGCCCCTC 220

Qy   852 CTTTCCCAAATTCTACACAACAGCTCAAAATCTGCAGTCTGAAACCCAAGCAGCCAGACA 911
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   219 CTTTCCCAAATTCTACACAACAGCTCAAAATCTGCAGTCTGAAACCCAAGCAGCCAGACA 160

Qy   912 CTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGCAGGAGTTGCAT 971
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   159 CTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGCAGGAGTTGCAT 100

Qy   972 CAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCCCACATTGTTGA 1031
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99 CAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCCCACATTGTTGA 40

Qy   357 CCCCCACCACGCGGTGAACACTATCTCCACCCGGTGCGGCTGGCATCGTCTGCACCCT 416
           ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   709 CCCCCACCACGCGCG-GAACACTATCTCCACCCGGTGCGGCTGGCATCGTCTGCACCCT 651

Qy   417 GTGGGCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTCGTGCA   476
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   650 GTGGGCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTCGTGCA   591

Qy   477 AGAGACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACAT 536
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   590 AGAGACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACAT 531

Qy   537 CATGTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGAT   596
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   530 CATGTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGAT   471

Qy   597 TGTTTGGAGCCTGAGGCGGAGGCGAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGAC 656
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   470 TGTTTGGAGCCTGAGGCGGAGGCGAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGAC 411

Qy   657 CCGGTCATCATGGTGGTGGCAATTGTGTCATCACATGCTACCTGCCCAGGCGTGTCTGC   716
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   410 CCGGTCATCATGGTGGTGGCAATTGTGTCATCACATGCTACCTGCCCAGCGTGTCTGC   351

Qy   717 TAGACTCTATTTCCTCTGGACGGTGCCTGCCGAGTCCTGCGATCCCTGTCCATGGGGGC   776
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   350 TAGACTCTATTTCCTCTGGACGGTGCCCTGCCGAGTGCCTGCGATCCCTGTCCATGGGGGC 291

Qy   777 CCTGCACATAACCCTCCAGCTTCACCTACACATGAACAGCATGCTCGGATCCCTGGTGTATTA 836
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   290 CCTGCACATAACCCTCCAGCTTCACCTACACATGAACAGCATGCTCGGATCCCTGGTGTATTA 231

Qy   837 TTTTTCAAGCCCCTCCTTCCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACC 896
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   230 TTTTTCAAGCCCCTCCTTCCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACC 171

Qy   897 CAAGCAGCCAGGACACTCAAAAACACAAAGGCCGGAGAGATGCCAATTTCGAACCTCG   956
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   170 CAAGCAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCG   111

Qy   957 TCGCAGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGA 1016
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   110 TCGCAGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGA 51

Qy   1017 TCCCCACATTGT 1028
            ||||||||||||
Db   50   TCCCCACATTGT 39

Qy   1032 GTGGGCAC 1038
            ||||||||
Db   39   GTGGGCAC 33

RESULT 5
CD559498/c
LOCUS   CD559498   786 bp   mRNA   linear   EST 19-NOV-2003
DEFINITION   AGENCOURT 14496820 NIH_MGC_195 Homo sapiens cDNA clone
             IMAGE:6971816 5', mRNA sequence.
ACCESSION   CD559498
VERSION   CD559498.2 GI:38453562
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 786)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   On Jun 10, 2003 this sequence version replaced gi:31585566.
             Contact: Daniela S. Gerhard, Ph.D.
             Office of Cancer Genomics
             National Cancer Institute / NIH
             Bldg. 31 Rm10A07 Bethesda, MD 20892
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: IRBK2 row:.c column: 07
             High quality sequence start:.3
             High quality sequence stop: 706.
FEATURES   Location/Qualifiers
     source   1..786
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6971816"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN

Query Match   65.1%; Score 676; DB 14; Length 786;
Best Local Similarity   98.0%; Pred. No. 8.1e-171;
Matches 717; Conservative   0; Mismatches 10; Indels   5; Gaps   3;

Qy   297 GGCCGGGAGCATCGTGTTCCTTACGGTGGCTGGCGACAGGTATTCAAAGTGTCCA 356
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   765 GGCCGGGAGCATCGTGT--CCTTACGGTGGCTGCGACAGGTATT--CAAAGTGTCCA 710

RESULT 6
CD559497/c
LOCUS   CD559497   793 bp   mRNA   linear   EST 19-NOV-2003
DEFINITION   AGENCOURT 14496881 NIH_MGC_195 Homo sapiens cDNA clone
             IMAGE:6971817 5', mRNA sequence.
ACCESSION   CD559497
VERSION   CD559497.2 GI:38453561
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 793)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   On Jun 10, 2003 this sequence version replaced gi:31585565.
             Contact: Daniela S. Gerhard, Ph.D.
             Office of Cancer Genomics
             National Cancer Institute / NIH
             Bldg. 31 Rm10A07 Bethesda, MD 20892
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Narayan Bhat

```
Qy   777  CCTGCACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCCTGGTGTGTATTA  836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   291  CCTGCACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCCTGGTGTGTATTA  232

Qy   837  TTTTTCAAGCCCCTCCTTCCTTCCCAAATTCTACACAGCTCAAAATCTGCAGTCTGAAACC  896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   231  TTTTTCAAGCCCCTCCTTCCTTCCCAAATTCTACACAGCTCAAAATCTGCAGTCTGAAACC  172

Qy   897  CAAGCAGCCAGGAGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGG  956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171  CAAGCAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGG  112

Qy   957  TCGCAGGAGTTGCATCAGTGTGGCCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGA  1016
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   111  TCGCAGGAGTTGCATCAGTGTGGCCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGA  52

Qy  1017  TCCCCACATTGT  1028
          ||||||||||||
Db    51  TCCCCACATTGT  40
```

RESULT 7
CD559496/c
LOCUS   CD559496   790 bp   mRNA   linear   EST 19-NOV-2003
DEFINITION AGENCOURT 14497010 NIH_MGC_195 5', mRNA sequence.
            IMAGE:6971819 5', mRNA sequence.
ACCESSION   CD559496
VERSION   CD559496.2   GI:38453560
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 790)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   On Jun 10, 2003 this sequence version replaced gi:31585564.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
            cDNA Library Preparation: Bhat Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK2 row: c column: 10
            High quality sequence stop: 703.
FEATURES   Location/Qualifiers
     source   1..790
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6971819"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC

            cDNA Library Preparation: Bhat Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK2 row: c column: 08
            High quality sequence start: 5
            High quality sequence stop: 715.
FEATURES   Location/Qualifiers
     source   1..793
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6971817"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."

ORIGIN

Query Match   64.8%; Score 672.8; DB 14; Length 793;
Best Local Similarity   97.7%; Pred. No. 5.9e-170;
Matches   715; Conservative   0; Mismatches   12; Indels   5; Gaps   3;

Qy   297 GGCCGGGAGCATCGTGTCCTTACGGTGGCTGCGCGGACAGGTATTCAAAGTGGTCCA 356
          |||||||||||||||||| | | ||||||||||||||||||||||||||||||||||||
Db   766 GGCCGGGAGCATCGTGT--CCTTACGGTGGCTGCGCGGACAG--TATTTCAAAGGGTCCA 711

Qy   357 CCCCCACCACGCGGTGAACACTATCTCACCCGGTGCGGCGGCTGGCATCGTCTGCCACCCT 416
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   710 CCCCCACCACGCGGG-GAACACTATCTCACCCGGTGCGGCGGCTGGCATCGTCTGCCACCCT 652

Qy   417 GTGGGCCCTGGTCATCCTCGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCA 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   651 GTGGGCCCTGGTCATCCTCGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCA 592

Qy   477 AGAGACGGCCGGTCCTCGTGGAGAGCTTCATCATGAGCAGTGGCGGCCAATGGCTGGCATGCAT 536
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   591 AGAGACGGCCGGTCCTCGTGGAGAGCTTCATCATGAGCAGTGGCGGCCAATGGCTGGCATGACAT 532

Qy   537 CATGTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGAT 596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   531 CATGTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGAT 472

Qy   597 TGTTTGGAGCCTGAGGCGGAGGCAGCAGCTCGGCCAGACAGGCTCGGATGAGAAGGCGAC 656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   471 TGTTTGGAGCCCTGAGGCGGAGGCAGCAGCTCGGCCAGACAGGCTCGGATGAGAAGGCGAC 412

Qy   657 CCGGTTCATCATGGTGGTGGCAATTGTGTCATCACATGCTACCTGCCCAGCCGGTCTGC 716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   411 CCGGTTCATCATGGTGGTGGCAATTGTGTTCATCATGCTACCTGCCCAGCCGGTCTGC 352

Qy   717 TAGACTTCATTTCCTCTGGACGCGTGCCCTCGAGTGCCTGCAGTCCCTCGTCTCATGGGGC 776
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   351 TAGACTTCATTTCCTCTGGACGCGTGCCCTCGAGTGCCTGCAGTCCCTCGTCTCATGGGGC 292
```

ORIGIN

Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presV.dat
a Note: this is a NIH_MGC Library."

Query Match          64.5%;     Score 669.4;  DB 14;  Length 790;
Best Local Similarity  97.6%;     Pred. No. 4.9e-169;
Matches  690;   Conservative   0;   Mismatches  16;   Indels  1;   Gaps  1;

```
Qy    328 GCTCACGGCAGGTATTTCAAAGTGGTCACCCCACCACGCGGTGAACACTATCTCCACC 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    735 GGTGCTCGCACAGGTATTCAAAGTGTTCACCCCCACCACGCGG-GAACACTATCTCCACC 677

Qy    388 CGGGTGGGCGGCTGCATCGTCTCGCACCCTGTGGGCCCTGGTCATCCTGGAACAGTGTAT 447
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    676 CGGGTGGGCGGCTGCATCGTCTCGCACCCTGTGGGCCCTGGTCATCCTGGAACAGTGTAT 617

Qy    448 CTTTTGCTGGAGAACCATCTCTGCGTGCAAGAGACGGCCGTCTCCTGTGAGAGCTTCATC 507
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616 CTTTTGCTAGAGAACCATCTCTGCGTGCAAGAGACGGCCGTCTCCTGTGAGAGCTTCATC 557

Qy    508 ATGGAGTCGGCCAATGGCTGGCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCTC 567
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    556 ATGGAGTCGGCCAATGGCTGGCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCCTC 497

Qy    568 GGCATCATCTTATTTTGCTCCTTCAAGATTGTTTGGAGCCTGAGGCGGCAGCAGTG 627
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 GGCATCATCATCTTATTTTGCTCCTTCAAGATTGTTTGGAGCCTGAGGCGGCAGCAGTG 437

Qy    628 GCCAGACAGGCTCGGATGAAGAAGGCCACCCGGTTCATCATGTGGTGGCAATTGTGTTC 687
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 GCCAGACAGGCTCGGATGAAGAAGGCGACCCGGTTCATCATGTGGTGGCAATTGTGTTC 377

Qy    688 ATCACATGCTACCTGCCCAGCGTGTCGCTAGACTCTATTTCCTCGACGGTGCCCTCG 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 ATCACATGCTACCTGCCCAGCGTGTCTGCTAGACTCTATTTCCTCGACGGTGCCCTCG 317

Qy    748 AGTGCCTGGATCCCTCTGTCCATGGGGCCCTGCACATAACCCTCAGCTTCACCTACATG 807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 AGTGCCTGGATCCCTCTGTCCATGGGGCCCTGCACATAACCCTCAGCTTCACCTACATG 257

Qy    808 AACAGCAGATGCTGGATCCCCTGGTGTATTATTTTCAAGCCCCTCCTTTCCCAAATTCTAC 867
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 AACAGCAGATGCTGGATCCCCTGGTGTATTATTTTCAAGCCCCTCCTTTCCCAAATTCTAC 197

Qy    868 AACAAGCTCAAAATCTGCAGTCTGAAACCCAGCAGCCAGGACACTCAAAAACACAAAGG 927
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 AACAAGCTCAAAATCTGCAGTCTGAAACCCAGCAGCCAGGACACTCAAAAACACAAAGG 137

Qy    928 CCGGAGAGATGCCAATTTCGAACCTCGGTCGCAGGAGTTGCATCAGTGTGGCAAATAGT 987
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 CCGGAGAGATGCCAATTTCGAACCTCGGTCGCAGGAGTTGCATCAGTGTGGCAAATAGT 77

Qy    988 TTCCAAAGCCAGTCAGTCTGATGGGATCCCCACATTGTTGAGTG 1034
          ||||||||||||||||||||||||||||||||||||||||||||
Db    76 TTCCAAAGCCAGTCAGTCTGATGGGATCCCCACATTGTTGAGGG 30
```

RESULT 8
CD559653
CD559653
LOCUS   CD559653   AGENCOURT 14496792 NIH_MGC_195   773 bp   mRNA   linear   EST   26-NOV-2003
DEFINITION IMAGE:6971816 5', mRNA sequence.
ACCESSION   CD559653
VERSION   CD559653.2   GI:38559026
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 773)

AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   On Jun 10, 2003 this sequence version replaced gi:31585721.
          Contact:Daniela S. Gerhard, Ph.D.
          Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Narayan Bhat
          cDNA Library Preparation: Bhat Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: IRBK2   row:c   column: 07
          High quality sequence start: 9
          High quality sequence stop: 740.

FEATURES   Location/Qualifiers
     source   1..773
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6971816"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note=Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presV.dat
                     a Note: this is a NIH_MGC Library."

ORIGIN

Query Match   59.7%;   Score 619.8; DB 14; Length 773;
Best Local Similarity 98.6%;   Pred. No. 1.2e-155;
Matches 624;   Conservative   0;   Mismatches 0;   Indels 9;   Gaps 0;

```
Qy    1 ATGACAACGGGTGCGTGCTGCCGCATCGAGGGGGGACACCATCTCCAGGTGATGCGCCG 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ATGTACAACGGGTGCGTGCTGCCGCATCGAGGGGGGACACCATCTCCAGGTGATGCGCCG 200

Qy    61 CTGCTCATGTGGCCTTTGTGCTGGGGCCGACTAGGCAATGGGAGTGGGTCGCGTGTGGTTTC 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CTGCTCATGTGGCCTTTGTGCTGGGGCCGACTAGGCAATGGGAGTGGGTCGCGTGTGGTTTC 260

Qy    121 TGCTTCCACATGAAGAGCTGGAAGCCCAGCACTCACTGTTCATTTGGCCTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 TGCTTCCACATGAAGAGCTGGAAGCCCAGCACTGTTCATTTTTACCTTTTCAATTTGGCCGTGGCT 320

Qy    181 GATTTCCCTTATGATCTGCCCTGCCTTTTCGGACAGACTATTACCTCAGCGTAGACAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 GATTTCCCTTATGATCTGCCCTGCCTTTTCGGACAGACTATTACCTCAGCGTAGACAC 380

Qy    241 TGGGCTTTTGGGACCATTCCCTGCGGAGTGGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 TGGGCTTTTGGGACCATTCCCTGCGGAGTGGGGGCTCTTCACGTTGGCCATGAACAGGGCC 440

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGCTGCCGACACGTATTCAAAGTGGTCCACCC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Left column

```
Db  441 GGGAGCATCGTGTTCCTTACGGTGGTGGTGGTCGCGGACAGGTATTTCAAAGTGGTCCACCCC 500
Qy  361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCCGCATCGTCTGGCCACCCTGTGG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  501 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCCGCATCGTCTGGCCACCCTGTGG 560
Qy  421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCCGTGCAAGAG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  561 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCCGTGCAAGAG 620
Qy  481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTGCGGCCAATGCTGGCATGACATCATG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  621 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTGCGGCCAATGCTGGCATGACATCATG 680
Qy  541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
        ||||||||||||||||||||||||| |||||||    |||||    ||||||
Db  681 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTTGCTCCTTCAGATTTGT 740
Qy  601 TGGAGCCTGAGGCGGAGGCGGCAGCAGCTGCCAGA 633
        |||||||||| || ||||||||||||||||||
Db  741 TGGAGCCTGANNGCGAAGCCAGCAGCAGCTGGCCAGA 773
```

RESULT 9
CD559650
LOCUS   CD559650   748 bp   mRNA   linear   EST 26-NOV-2003
DEFINITION AGENCOURT_14497045 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:6971820 5', mRNA sequence.
ACCESSION   CD559650
VERSION   CD559650.2   GI:38559022
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 748)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   On Jun 10, 2003 this sequence version replaced gi:31585718.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
            cDNA Library Preparation: Bhat Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK2 row: c column: 11
            High quality sequence start: 11
            High quality sequence stop: 745.
FEATURES   Location/Qualifiers
   source   1..748
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6971820"
            /tissue_type="mixed"
            /lab_host="DH5A (T1 phage-resistant)"
            /clone_lib="NIH_MGC_195"
            /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
            loxP-HindIII; Clones from this library have been
            PCR-amplified using gene-specific primers to contain the
            complete open reading frame (based on known gene sequences
            available from NCBI's RefSeq). Template for PCR is cDNA
            derived from either pooled cytoplasmic polyA RNA from 30
            cells lines or pooled total RNA from 10 different tissues
            (from BD Biosciences/Clontech and Washington University).
            PCR products are directionally cloned into the loxP sites
```

Right column

```
            of the pDNR-Dual vector. Library constructed by Dr.
            Narayan Bhat, Earl Bere III and Hongling Liao (Gene
            Expression Laboratory, Research Technology Program, SAIC
            Frederick, NCI-Frederick, Frederick, MD 21702). For
            information on which gene each clone represents, please
            visit our anonymous ftp site at
            ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presV.dat
            a Note: this is a NIH_MGC Library."
ORIGIN
Query Match         57.8%;   Score 600;  DB 14;   Length 748;
Best Local Similarity  99.2%;  Pred. No. 2.5e-150;
Matches  603;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCAGGTGATGCCGCCG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCAGGTGATGCCGCCG 200
Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGCCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 CTGCTCATTGTGGCCTTTGTGCTGGCCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 260
Qy  121 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTCAATTGGCCGTGCTGCT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  261 TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTCAATTGGCCGTGCTGCT 320
Qy  181 GATTTCCTTATGATCTGCCTGCCTTTCGGACAGACATTACCTCAGACGTAGACAC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  321 GATTTCCTTATGATCTGCCTGCCTTTCGGACAGACATTACCTCAGACGTAGACAC 380
Qy  241 TGGGCTTTGGGGACATTCCTGCCGAGTGGGCTCTCACGTGGCATGAACAGGGCC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  381 TGGGCTTTGGGGACATTCCTGCCGAGTGGGCTCTCACGTGGCATGAACAGGGCC 440
Qy  301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGTCCACCCC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  441 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAAGTATTTCAAAGGCGTCCACCCC 500
Qy  361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCCATCGTCTGGCCACCCTGTGG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  501 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCCATCGTCTGGCCACCCTGTGG 560
Qy  421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCCGTGCAAGAG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  561 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCCGTGCAAGAG 620
Qy  481 ACGGCCGTCTCCGTGAGAGCTTCATCATGGCCAATGGCCAATGGCCATGACATCATG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  621 ACGGCCGTCTCCGTGAGAGCTTCATCATGAGTGGCCAATGGCCATGACATCATG 680
Qy  541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  681 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTTGCTCCTTCAGATTGTT 740
Qy  601 TGGAGCCT 608
        ||||||||
Db  741 TGGAGCCT 748
```

RESULT 10
CD559651
LOCUS   CD559651   743 bp   mRNA   linear   EST 26-NOV-2003
DEFINITION AGENCOURT_14496981 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:6971819 5', mRNA sequence.
ACCESSION   CD559651
VERSION   CD559651.2   GI:38559023
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 743)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.

```
Qy   361  CACCACGGGTGAACACTATCTCCACCCGGTGCGGCTGGCATCGTCTGCACCCTGTGG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   471  CACCACGGGTGAACACTATCTCCACCCGGTGCGGCTGGCATCGTCTGCACCCTGTGG  530

Qy   421  GCCCTGGTCATCCTGGGAACACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   531  GCCCTGGTCATCCTGGGAACACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  590

Qy   481  ACGGCCGTCCTCTGAGAGCTTCATCATGGAGTCGGCCAATGAGTCTGGCATGACATCATG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   591  ACGGCCGTCCTCTGAGAGCTTCATCATGGAGTCGGCCAATGAGTCTGGCATGACATCATG  650

Qy   541  TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTGCTCCTTCAAGATTGTT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   651  TTCCAGCTGGAGTTCTTTATGCCCCTGGC-TCATCTTATTTGCTCCTTC-AGATTGTT  708

Qy   601  TGGAGCCTGAGGCGGAG  617
          |||||||||||||||||
Db   709  TGGAGCCTGAGGCGGAG  725
```

RESULT 11

```
CD559652
LOCUS       CD559652             731 bp    mRNA    linear    EST 26-NOV-2003
DEFINITION  AGENCOURT_14496854 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:6971817 5', mRNA sequence.
ACCESSION   CD559652
VERSION     CD559652.2  GI:38559024
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Jun 10, 2003 this sequence version replaced gi:31585720.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
            cDNA Library Preparation: Bhat Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK2  row: c  column: 08
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1..731
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6971817"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH MGC 195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
```

(Left column continued)

```
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     On Jun 10, 2003 this sequence version replaced gi:31585719.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
            cDNA Library Preparation: Bhat Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK2  row: c  column: 10
            High quality sequence stop: 728.
FEATURES             Location/Qualifiers
     source          1..743
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6971819"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH MGC 195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSv.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN

Query Match    57.0%;  Score 591.4;  DB 14;  Length 743;
Best Local Similarity  99.5%;  Pred. No. 5.1e-148;
Matches  614;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  2;

Qy     1  ATGTACAACGGGTCGTCGCCGCATCGAGGGGGACACCATCTCCAGGTGATGCCGCCG  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   111  ATGTACACAGGGTCGTCGCCGCGTCGCCGCATCGAGGGGGACACACCATCTCCCAGGTGATGCCGCCG  170

Qy    61  CTGCTCATTGTGCCTTTGTGCTGGGGCACCAGGCCAATGGGGGTCGCCCGTGTGGTTTC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171  CTGCTCATTGTGGCCTTTGTGCTGGGGCACCAGGCCAATGGGGGTCGCCCGTGTGGTTTC  230

Qy   121  TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTGGCCGTGGCT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   231  TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTGGCCGTGGCT  290

Qy   181  GATTCCTCCTTATGATCTGCCTGCCTTTTCGACAGACTATTACCTCAGACGTAGACAC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   291  GATTTCCTCCTTATGATCTGCCTGCCTTTTCGACAGACTATTACCTCAGACGTAGACAC  350

Qy   241  TGGGGCTTTTGGGGACATTCCCTGCCGAGTGGGCGCTCTTCACGTTGGCCATGAACAGGGCC  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   351  TGGGGCTTTTGGGGACATTCCCTGCCGAGTGGGCGCTCTTCACGTTGGCCATGAACAGGGCC  410

Qy   301  GGGAGCATGTGTTCCTTACGGTGGTGGTGCGCGACAGGTATTTCAAAGTGGTCCACCCC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   411  GGGAGCATCGTGTTCCTTACGGTGGTGGTGCGCGACAGGTATTTCAAAGTGGTCCACCCC  470
```

ORIGIN

Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presV.dat
a Note: this is a NIH_MGC Library."

Query Match   56.3%;   Score 584.4; DB 14; Length 731;
Best Local Similarity 97.0%;   Pred. No. 3.9e-146;
Matches 606;   Conservative 0; Mismatches 17; Indels 2; Gaps 1;

Qy   1 ATGTACAACGGGTGCGTGCTGCCGCATGCGAGGGGACACCATCTCCCAGGTGATGCCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   92 ATGTACAACGGGTCGTGCTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 151

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   152 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 211

Qy   121 TGCTTCCACATGAAGAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212 TGCTTCCACATGAAGAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 271

Qy   181 GATTTCCTCCTTATGATCGCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   272 GATTTCCTCCTTATGATCGCCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 331

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   332 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGGCTCTTCACGTTGGCCATGAACAGGGCC 391

Qy   301 GGGAGCCATCGTGTTCCTTACGGTGGTGCTGCCGGACAGGTATTTCAAAGTGGTCCACCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   392 GGGAGCCATCGTGTTCCTTACGGTGGTGCTGCCGGACAGGTATTTCAAAGTGGTCCACCCC 451

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   452 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG 511

Qy   421 GCCCTGGTCATCCTGTCTGTGAGAGTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   512 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCCGCGGTGCAAGAG 571

Qy   481 ACGGGCCGTCCTGTGAGAGTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   572 ACGGGCCGTCCTGTGAGAGTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 631

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
          |||||||||||||||||| ||||||||||
Db   632 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCATCTTATTT--GCTCCTTCAGATTGTT 689

Qy   601 TGGAGCCCAGGCGGAGGCAGCAGC 625
          ||||||| ||||||||||
Db   690 TGGAGCTGANGCGGAGCACAGCTGC 714

RESULT 12
CD246157
LOCUS   CD246157   1004 bp   mRNA   linear   EST 22-MAY-2003
DEFINITION   AGENCOURT_14127000 NIH_MGC_145 Homo sapiens cDNA clone
             IMAGE:6912801 5', mRNA sequence.
ACCESSION   CD246157
VERSION   CD246157.1 GI:31006621
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1004)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.

FEATURES   Location/Qualifiers
     source   1..1004
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6912801"
                     /tissue_type="mixed"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_145"
                     /note="Vector: pcDNA3.1; Site 1: varies by clone; Site 2:
                     varies by clone; ORFs were PCR-amplified and cloned into
                     pcDNA3.1 by the GPCR Consortium. Cloning sites vary by
                     clone and include the following: 5'-EcoRV-XmnI/XhoI-3',
                     5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional).
                     For information about which gene each clones represents,
                     please visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBI.presV.dat
                     a Note: this is a NIH_MGC Library."

ORIGIN

Query Match   53.8%;   Score 558.6; DB 14; Length 1004;
Best Local Similarity 99.3%;   Pred. No. 4.2e-139;
Matches 561;   Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTGCGTGCTGCCGCATCGAGGGGGACACCATCTCCAGGTGATGCCGCCG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 190

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTATGGTTTC 250

Qy   121 TGCTTCCACATGAAGAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTTGGCCGTGGCT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 TGCTTCCACATGAAGAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTTGGCCGTGGCT 310

Qy   181 GATTTCCTCCTTATGATCTGCTGCCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 GATTTCCTCCTTATGATCTGCCTGCCCTTTTCGGACAGACTATTACCTCATACGTAGACAC 370

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGGCTCTTCACGTTGGCCATGAACAGGGCC 430

Qy   301 GGGAGCCATCGTGTTCCTTACGGTGGTGCTGCCGGACAGGTATTTCAAAGTGGTCCACCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 GGGAGCCATCGTGTTCCTTACGGTGGTGCTGCCGGACAGGTATTTCAAAGTGGTCCACCCC 490

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG 550

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCCGCGGTGCAAGAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   551 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCCGCGGTGCAAGAG 610

Qy   481 ACGGGCCGTCCTGTCCGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   611 ACGGGCCGTCCTGTCCGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 670

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCC 565
          |||||| ||||| ||||

Db   671 TTCCAACTGGAGTTCTTTATACCCC 695

RESULT 13
CB432388/c
LOCUS   CE432388   594 bp   DNA   linear   GSS 27-SEP-2003
DEFINITION   tigr-gss-dog-1700036299607 Dog Library Canis familiaris genomic,
             genomic survey sequence.
ACCESSION   CE432388
VERSION   CE432388.1 GI:36707139
KEYWORDS   GSS.
SOURCE   Canis familiaris (dog)
ORGANISM   Canis familiaris
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 594)
AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
             Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
             Venter,J.C.
TITLE   The dog genome: survey sequencing and comparative analysis
JOURNAL   Science 301 (5641), 1898-1903 (2003)
MEDLINE   22875432
PUBMED   14512627
COMMENT   Contact: Kirkness EF
             The Institute for Genomic Research
             Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
             Rockville, MD 20850, USA
             Tel: 301-838-0200
             Fax: 301-838-0208
             Email: ekirknes@tigr.org
             Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..594
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site 1: BstXI; Libraries were prepared from
                     peripheral blood"

ORIGIN

Query Match   43.7%; Score 453.8; DB 29; Length 594;
Best Local Similarity 85.3%; Pred. No. 5.8e-111;
Matches 506; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy   376 ACTATCTCCACCCGGGTGGCCGGCTGGCATCGTCTCGCACCCTGGGGCCCTGGTCATCCTG 435
              ||||||||||||||||||||||| ||||| || |||||||| |||||||||||||||||||
Db   594 ACCATTTCCAACCGGACTGCAGCTCAGCTGGCATCGTCGTGTGCCCTTTGGATCATGGTCATCCTG 535

Qy   436 GGAACAGTGTATCTTTTTGCTGAGAACCATCTCGCGTGCAAGAGACGGCCGCGTCCTCCTGT 495
              |||| ||| ||||||| |||| || |||||||||||| | || ||| |||||||||||||
Db   534 GGCACTCTACACCTTTTGATGAGGAACCATCTGTGCGCATGAGAAGACCATATCTTGT 475

Qy   496 GAGAGCTTCATCATGAGAGTCGGCCAATGACATGCAACATCATGTTCCAGCTGGAGTTC 555
              |||||||||||||||||| |||||||||||| |||||||||||||||||||||| ||||
Db   474 GAGAGCCTCATCATGAGAGTCAGCCAATGGCTGGCACCGCACCACATCATGTTCCAGCTGGAATTC 415

Qy   556 TTTATGCCCCCTTCGGCATCCATCATCTATTTTGCTCCTTCAAGATTGTTGGAGCCTGAGCGGG 615
              |||| ||||| ||||| ||| |||||| |||||||||||||||| |||| ||||||||||||
Db   414 TTCCTCCCTCTGGCATCCATCATCCGTGTTCTGCTCCTTCAGGATTAATTGGAGTCTGAAGCAG 355

Qy   616 AGGCAGCAGTCGGCCAGAGCAGGCTCGGATGAGAAGAGGCGACCCGGTTCATCATGGTGGTG 675
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   354 AGGCGGCAGTCGGCCAGGCAGACTCGGATGAGAAGAAGGCTACCCGGTTCATCATGGTGTGTG 295

Qy   676 GCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGACTCTATTTCCTTGG 735
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||
Db   294 GCGGTTGTGTTCATCACGTGCTACCTGCCCCAGCGTCGGCCAGACTCTTATTTCCTCTGG 235

Qy   736 ACGGTGCCCCTCGAGTGCCTTGCCTCGCTGCCTGCTGGGCCTCCATGGGGCCTGCACATAACCCTCAGC 795
              |||||||||||||||||||||||||||||||||||||| || |||||||||||||||||| ||||
Db   234 ACGGTGCCCTCGAGTGCCTTGCCTGGACCCCTGTCCACGTAGCCCTCACGTAGCCCTCAGC 175

Qy   796 TTCACCTACACTGAACAGCATGCTGGATCCCTCGTGTGTATATTTTTCAAGCCCCTCCTTT 855
              |||||||||||||||||||||||||||||||||||| |||| |||| ||| ||||||||||
Db   174 TTCACCTACACTGAACAGCATGCTGGATCCCTCGTGTGTATTATTTTTCGAGTCCTGTATTC 115

Qy   856 CCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAAACCCAAGCAGCCAGCAGACACTCA 915
              |||||||||||||||||||||||||| ||||||||||| || |||||||||||||||||||
Db   114 CCCAAATTCTACACCAAGCTCAAGATCGCGGTTTGCGACCAAAGAGTCCAGGGCACTCC 55

Qy   916 AAAACACAAAGGCCGGAAGAGATGCCAATTTGCAAATTCGAACCTCGGTCGCAGGAGTTG 968
              |||| |||||||| | ||||||||||| |||| |||||||| || |||||||||||||||
Db   54 AAGACCCAGAGCCCGAAGACAGATGCCAATCCCAAAGCTCTGTCGCAAGAGTTG 2

RESULT 14
CB576966
LOCUS   CB576966   635 bp   mRNA   linear   EST 03-APR-2003
DEFINITION   AMGNNUC:URGP1-00001-C7-A urgp1 (14349) Rattus norvegicus cDNA clone
             urgp1-00001-c7 5', mRNA sequence.
ACCESSION   CB576966
VERSION   CB576966.1 GI:29521007
KEYWORDS   EST.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 635)
AUTHORS   Amgen EST Program.
TITLE   Amgen Rat EST Program
JOURNAL   Unpublished (2003)
COMMENT   Contact: Dan Fitzpatrick
             Amgen, Inc
             One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
             Tel: 805 447-4881
             Plate: 00001 row: c column: 7.
FEATURES   Location/Qualifiers
     source   1..635
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="urgp1-00001-c7"
                     /clone_lib="urgp1 (14349)"
                     /note="Vector: pSPORT1; Rat GPCR library rearrayed
                     internal pSPORT vector"

ORIGIN

Query Match   42.8%; Score 443.8; DB 14; Length 635;
Best Local Similarity 82.0%; Pred. No. 3e-108;
Matches 511; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy   5 ACAACGGGTCGTGCTGCGGCGGGGACACCATCTCCCAGTGATGGGTGATGGCCGGCTGC 64
              |||||||||||||||||||||||||||| |||||||||||||| || ||| ||||||||
Db   12 ACAACGGGTCGTGCTGTCTCATCGAGGGGAACCCATCACCCAGTAATGCCACCTTTAC 71

Qy   65 TCATTGTGGCCTTTGTGCTGCGGGCGCACTAGGCAATGGGTCGCCCTGTGGTTCTCGCT 124
              |||||||||||| || |||| ||||||||||||||||||||||| |||||||||||||||
Db   72 TCATCCTGCCCTTTCCTGCTGGAGCCCTGGGCAACGGCCTAGCCCTGTGTGGTTCTCGCT 131

Qy   125 TCCAATCAGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTTGGGCCGTGGCTGATT 184
              |||| ||||||||||||| |||||||||| |||||||||||| |||||||||||||||||
Db   132 TTCACATCAGAAGACCTGGAAGTCGAAGTTGGCACTATTTACCTTTCAACTGGCTAGCCGATT 191

Qy   185 TCCTCCTTATGATCTGCCTGCCTTTCGGACGACTATTACCTCAGACTAGACACTCGGG 244
              ||||| |||| ||||||||||| ||||||||||||||||||||||||||||||||||||
Db   192 TTCTCCCTCATGATCTGCCTGCTCTTCACGTTGGCCATGAACTCTGCGGCCATGATTGGA 251

Qy   245 CTTTTGGGACATTCCCTGCCAGTGGGGCGTCTTCACGTTGGCCATGAACACGGCCGGGA 304
              ||||||||||| |||||| || |||| ||||||||||||||||||||||||||||||||
Db   252 TTTTGGGGGATATTCCCTGCCGGCGCCGGCTGTCTCTTCTCATCGTCGGCCATGAATGAGGCCGGAA 311

Qy   305 GCATCGTGTTCCTTACGGTGGTGCTGCGCGGACAGGTATTTCAAAGTGGTGCACCCCCACC 364
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   305 GCATCGTGTTCCTTACGGTGGTGCTGCGGACAGGTATTTCAAAGTGGTGCATCCCCACC 364

```
                                                    /clone="IMAGE:5222688"
                                                    /tissue_type="Pancreas, Spleen, adult pooled"
                                                    /clone_lib="NIH_MGC_120"
                                                    /lab_host="DH10B"
                                                    /note="Vector: pCMV-SPORT6"

ORIGIN

        Query Match      35.7%;   Score 370.8;  DB 11;  Length 2146;
        Best Local Similarity   64.1%;   Pred. No. 2.8e-88;
        Matches  599;  Conservative   0;  Mismatches  317;  Indels  18;  Gaps  2;

Qy      1 ATGTACAACGGGTCGTGCGTGCCGATCGAGGGGGACACCATCTCCAGTGATGCCGCCG  60
            ||||| |||| |  |  ||||||||||  |   ||  ||  |||  |||||| |||||
Db     97 ATAGACAAGAAGAACTGCGCGGTGTTCCGAGATGACTTCATTGTCAAGGTGTTGCCGCCG  156

Qy     61 CTGCTCATTGTGGCCTTTGTCGTGGCGGACTAGGCAATGGGGTCGCCCTGTGTGGTTTC  120
            ||| | |||  | |  ||  || |   || |||  |   | ||||| |||||||||||
Db    157 GTGTTGGGGCTGGAGTTTATCTTCGGGCTTCTGGGCCAATGCCTGCCCTGGTGGATTTTC  216

Qy    121 TGTTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTTGGCCGTGGCT  180
            |||||| |||  | ||    | |||| ||| | |  ||||| |||  | ||| |||||
Db    217 TGTTTCCACCTCAAGTCCTGGAAATCAGCCGGATTTCCTGGTTCAACCTGGCAGTGGGCT  276

Qy    181 GATTTCCTCCTTATGATCTGCCTCGGACAGCAGTATTACCTCAGACGTAGTAGACAC  240
            |||| |||  |  || || |||  ||||| ||| ||| |  || ||| |||| ||||
Db    277 GACTTTCTACTGATCATCTGCCTGCCCCTTCCTGATGGACAACTATGTGAGGCGTTGGGAC  336

Qy    241 TGGGCTTTTGGGGACCATTCCCTGCCGAGTGGGGGGGCTTCTTCACGTTGGCCATGAACAGGCC  300
            ||| | |  ||||| |||  |  | ||  | |  ||| ||  | ||| |||| ||| |||||
Db    337 TGGAAGTTTGGGGACATCCCTTGCCGGCTCATGGCTTCATGTGCTATGGAGTTGGCATGCCCAG  396

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGCGGACAGGTATTTCAAAGTGGTCACCCC  360
            ||| | |||  | | ||  ||  | || ||| | || || |||| |||  ||||
Db    397 GGCAGCATCATCTTCCTCACGGTGGTGGCCGTAGACAGGTATTTCCGGGTGGTGGTCCATCCC  456

Qy    361 CACCACGCGGGTGAACACTATCTCCACCACCGGGTGCGCGGTGGCTGGCTGCACCCTGTGG  420
            ||| ||||  ||||     ||    |  |  | |     |||||   || |||  ||||
Db    457 CACCAGCGCCCTGAACAAGATCTCCAATCGGACAGCAGCCATCATCTTCTGCGCTCTGTGG  516

Qy    421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG  480
            |||  | |  ||||| |||| ||  |  ||  |   ||  ||   | ||  ||||| |
Db    517 GGCATCACTATTGGCTGACAGTCCACCTCCTGAAGAGAAGATGCCGATCCAGAATGGC  576

Qy    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG  540
            ||||| |  | |  || | ||  || ||| | |  || ||||| ||| | | |||||
Db    577 GGTGCAAATTGTGCAGCAGCTTCAGCATCTGCCTATACCTTCCAGTGGCACGAAGCCATG  636

Qy    541 TTCCAGGTGGAGTTCTTTATGCCCCTGGCACATCTTATTTGCTCCTTCAAGATTGTT  600
            |||| | |   | || || |||  || || | |  ||||| | |  || ||| ||
Db    637 TTCCTCCTGGAGTTCTTCCTGCTGGCATCATCCTGTTGTCGTCTGCTCAGCCAGAATTATC  696

Qy    601 TGGAGCCCTGAGGCCGGAGGCCAGCAGCTGGGCTGCCAGGTGCCCTGCTCGACGGTGCCCTC-------GAGTGGCTGCTGGATCCCTCT  765
            |||||||||| | | |   | |  ||||| |      ||  ||||   ||| | ||| |||       |||| ||||| || |||||
Db    697 TGGAGCCTGCGCGGCCAGAG--ACAAATGGACCGGCATGCCCATGCCAGATCAAGAGACCATCACC  753

Qy    661 TTCATCATGGTGGTGGCAATTGTGTCATCACATGCTACCTGCCAGCCGTGTCTGCTAGA  720
            ||||| || |  ||||| |||  |  |  | |||  || ||  |||||  ||  ||  ||
Db    754 TTCATCATGTTGGTGGTGGCCATCATCGTCTTGTCATCGTCTTGTATCTGACTTCAGCCCAGGTGTTGGGG  813

Qy    721 CTCTATTTCCTCTGCGACGTGCCCTC--------GAGTGGCTGCTGGATCCCTCT  765
            || |||||| | |    |||| |||        |||| ||||| || |||||
Db    814 ATCCGCATCTTCTGCCTCCTGCACACTTCGGGCACGCAGAATTGTGAAGTGTACCGCTCG  873

Qy    766 GTCCATGGGCCCCTGCACATAACCCTCAGCTTCACCATGAACAGCATGCTGGATCCC  825
            || || |  | ||| ||| |  || ||||||| |    ||||   ||   |||  ||
Db    874 GTGGACCTGGCGTTCTTTATCACTCTCAGCTTCACCTACATGACAACAGCATGCTGGACCCC  933

Qy    826 CTGGTGTATTATTTTCCAGCCCCCCTTCCCAAATTCTACAACAAGCTCAAAATCTGC  885
            |||| |||  || | | |||  | |||  | ||||| | ||| | | ||| | |||
Db    934 GTGGTGTACTCTTCCCAGCCCATCCTTTCCCAACTTCTTCCCACTTGGATCAACCGC  993

Qy    886 AGTCTGAAACCCAGCAGCCAGCAGGACACTCAAAAA  919
            || |||  |  | |
```

```
Db    312 GCAATGTCTTCCTCACTGTGGTGGCCCTGGACAGGTATTTCAAAGTGGTCACCCCCACC  371
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    365 ACGGCGGTGAACACTATCTCCACCCACGGGGTGGCGGCTGGCCATCGTCTGCACCCTGTGGGCCC  424
            ||  |||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    372 ATATGGTGAACGCCATCTCCAATCGGACTGGAGCTCAGCTGCGTCTGTGTCCTCTGGACTT  431
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    425 TGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAGACGG  484
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 TGGTCATCCTGGGACTGTGTATCTTCTGATGAGAGACTCACCTGTGTGTGCGGGGGATGG  491
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    485 CCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATGTCC  544
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 TGTCATCTTGTGAGAGCTTCATCATGGAGCCCAACGGGGTGGCACGATATCATGTTCC  551
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    545 AGCTGGAGTTCTTTATGCCCCTCGGCATCATCATCTTATTTGCTCCTTCAAGATTGTTTGGA  604
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 AGCTGGAGTTCTTCCTGCCCCTGACCATCTGTTCTGTTCCTGCCCTCCTCAAAGTTGTTTGGA  611
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    605 GCCTGAGGCGGAGGCAGCAGCTG  627
            ||||||||||||||||||||||
Db    612 GCCTGAGCACAGAGGCAACAGCTG  634
```

RESULT 15
BC027965
LOCUS   BC027965   2146 bp   mRNA   linear   HTC 01-MAY-2002
DEFINITION   Homo sapiens, Similar to putative chemokine receptor; GTP-binding
             protein, clone IMAGE:5222688, mRNA.
ACCESSION   BC027965
VERSION   BC027965.1 GI:20379752
KEYWORDS   HTC.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2146)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-APR-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
             Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: National Institutes of Health Intramural
             Sequencing Center (NISC)
             Gaithersburg, Maryland;
             Web site: http://www.nisc.nih.gov/
             Contact: nisc_mgc@nhgri.nih.gov
COMMENT   Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
             Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
             Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
             Hansen,N., Ho,S.-L., Karlins,E., Laric,P., Legaspi,R., Maduro,Q.L.,
             Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
             Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C.,
             Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A.,
             Zhang,L.-H. and Green,E.D.
             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAK Plate: 49 Row: 1 Column: 24
             This clone was selected for full length sequencing because it
             passed the following selection criteria: matched mRNA gi: 5174460
             This clone has the following problem: frame shifted.
FEATURES   Location/Qualifiers
     source   1..2146
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"

CC   modulating cellular polypeptide expression or activity, useful as
CC   antagonists and agonists in disease treatment
XX
SQ   Sequence 346 AA;

Query Match   100.0%;   Score 1853;   DB 4;   Length 346;
Best Local Similarity   100.0%;   Pred. No. 5.9e-199;
Matches 346;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60

Qy   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Qy 121 HHAVNTISTRVAAGIVCLWALVILGTVYILLLENHLCVQETAVSCESFIMESANGWHDIM 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 HHAVNTISTRVAAGIVCLWALVILGTVYILLLENHLCVQETAVSCESFIMESANGWHDIM 180

Qy 181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240

Qy 241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300

Qy 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346

RESULT 3
AAU06197
ID   AAU06197 standard; protein; 346 AA.
XX
AC   AAU06197;
XX
DT   19-DEC-2001 (first entry)
XX
DE   Novel human G protein-coupled receptor (GPCR) protein.
XX
KW   Human; G-protein coupled receptor; GPCR; chemokine receptor; protease;
KW   hyperproliferative disorder; neurological disorder; psychiatric disease;
KW   inflammatory disorder; respiratory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200173029-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US009522.
XX
PR   27-MAR-2000; 2000US-0192419P.
PR   06-SEP-2000; 2000US-0230459P.
PR   20-SEP-2000; 2000US-0066535.
XX
PA   (PEKE) PE CORP NY.
XX
PI   Ye J, Cravchik A, Di Francesco V, Beasley EM;
XX
DR   WPI; 2001-616503/71.
DR   N-PSDB; AAS12581.
XX
PT   Novel human G-protein coupled receptor proteins and nucleic acid
PT   molecules encoding the protein for use in developing human therapeutics
PT   and diagnostic compositions and for identifying modulators of the
PT   protein.
XX
PS   Claim 1; Fig 1; 66pp; English.

XX
CC   The present invention relates to the isolation of a novel human G-protein
CC   coupled receptor (GPCR) which is related to the chemokine receptor
CC   subfamily. The cDNA and gene sequences encoding for GPCR are also given
CC   in the invention. The sequences of the invention are useful for
CC   diagnosing and treating diseases or conditions mediated by human
CC   proteases. Such diseases include hyperproliferative disorders (e.g.
CC   hyperplasia), neurological disorders (e.g. Parkinson's disease),
CC   psychiatric diseases (e.g. schizophrenia), inflammatory disorders (e.g.
CC   diabetes) and respiratory disorders (e.g. adult respiratory distress
CC   syndrome, ARDS). The GPCR protein is also useful for identifying a
CC   modulator of the expression of the protein. It also serves as a target
CC   for identifying agents for use in mammalian therapeutic applications,
CC   e.g. a human drug, particularly modulating a biological or pathological
CC   response in a cell or tissue that expresses the protein, in biological
CC   assays related to GPCRs that are related to members of the chemokine
CC   receptor subfamily, in drug screening assays and in competition binding
CC   assays. GPCR is also useful in diagnosing a disease or predisposition to
CC   a disease mediated by the peptide, in pharmacogenomic analysis. The
CC   polynucleotide sequences can also be used in gene therapy. The present
CC   sequence represents the novel human GPCR of the invention
XX
SQ   Sequence 346 AA;

Query Match   100.0%;   Score 1853;   DB 4;   Length 346;
Best Local Similarity   100.0%;   Pred. No. 5.9e-199;
Matches 346;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60

Qy   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Qy 121 HHAVNTISTRVAAGIVCLWALVILGTVYILLLENHLCVQETAVSCESFIMESANGWHDIM 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 HHAVNTISTRVAAGIVCLWALVILGTVYILLLENHLCVQETAVSCESFIMESANGWHDIM 180

Qy 181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240

Qy 241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300

Qy 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346

RESULT 4
AAU04373
ID   AAU04373 standard; protein; 346 AA.
XX
AC   AAU04373;
XX
DT   23-OCT-2001 (first entry)
XX
DE   Human G-protein coupled receptor, hRUP19.
XX
KW   Human; G-protein coupled receptor; GPCR; hRUP19; agonist;
KW   inverse agonist; lung cancer.
XX
OS   Homo sapiens.
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US031509.

```
Qy      301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
             |||||||||||||||||||||||||||||||||||||||||||||
Db      301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346

RESULT 5
AAB16172
ID   AAB16172 standard; protein; 346 AA.
XX
AC   AAB16172;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human G-protein coupled receptor 3 (GCREC-3) protein.
XX
KW   Human; G-protein coupled receptor 3; cell proliferative disorder;
KW   arteriosclerosis; hepatitis; cancer; neurological disorder; epilepsy;
KW   Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW   atherosclerosis; hypertension; myocardial infarction; peptic ulcer;
KW   gastrointestinal disorder; dysphagia; anorexia; autoimmune disorder;
KW   acquired immune deficiency syndrome; inflammatory disorder; infection;
KW   Addison's disease; allergy; Grave's disease; metabolic disorder; AIDS;
KW   diabetes; obesity; osteoporosis; gene therapy; GCREC-3.
XX
OS   Homo sapiens.
XX
FH   Key          Location/Qualifiers
FT   Peptide      1..32
FT                /label= Signal_peptide
FT   Domain       20..44
FT                /note= "Transmembrane domain"
FT   Protein      33..346
FT                /label= Human_mature_GCREC-3_protein
FT   Domain       93..110
FT                /note= "Transmembrane domain"
FT   Domain       137..154
FT                /note= "Transmembrane domain"
FT   Domain       222..244
FT                /note= "Transmembrane domain"
XX
PN   WO200187937-A2.
XX
PD   22-NOV-2001.
XX
PF   17-MAY-2001; 2001WO-US016285.
XX
PR   18-MAY-2000; 2000US-0205628P.
PR   22-MAY-2000; 2000US-0206222P.
PR   25-MAY-2000; 2000US-0207566P.
PR   02-JUN-2000; 2000US-0208834P.
PR   02-JUN-2000; 2000US-0208861P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Patterson C, Lu DAM, Thornton M, Lu Y, Tribouley CM, Graul R;
PI   Khan FA, Gandhi AR, Walia NK, Nguyen DB, Yue H, Hafalia A;
PI   Elliott VS, Lal P, Reddy R, Kallick DA, Tang TY, Au-Young J;
XX
DR   WPI; 2002-089844/12.
DR   N-PSDB; AAD26371.
XX
PT   Novel G-protein coupled receptors and polynucleotides useful for
PT   diagnosis, treatment and prevention of disorders of cell proliferation,
PT   neurological, cardiovascular, metabolic disorders and viral infections.
XX
PS   Claim 1; Page 105-106; 115pp; English.
XX
CC   The invention relates to human G-protein coupled receptor (GCREC)
CC   polypeptides and polynucleotides. GCREC polypeptides are useful for
CC   screening compounds that modulate their activity. They are useful in the
CC   diagnosis, prevention and treatment of disorders which include cell
CC   proliferative disorders such as arteriosclerosis, hepatitis,
```

```
XX
PR   17-NOV-1999;   99US-0166088P.
PR   17-NOV-1999;   99US-0166099P.
PR   17-NOV-1999;   99US-0166169P.
PR   23-DEC-1999;   99US-0171900P.
PR   23-DEC-1999;   99US-0171901P.
PR   11-FEB-2000;   99US-0171902P.
PR   14-MAR-2000;   2000US-0181749P.
PR   14-MAR-2000;   2000US-0189258P.
PR   10-APR-2000;   2000US-0189259P.
PR   10-APR-2000;   2000US-0195898P.
PR   28-APR-2000;   2000US-0195899P.
PR   12-MAY-2000;   2000US-0196078P.
PR   12-JUN-2000;   2000US-0200419P.
PR   12-JUN-2000;   2000US-0203630P.
PR   21-AUG-2000;   2000US-0210741P.
PR   26-SEP-2000;   2000US-0210982P.
PR   26-SEP-2000;   2000US-0226760P.
PR   20-OCT-2000;   2000US-0235418P.
PR   20-OCT-2000;   2000US-0235779P.
PR   24-OCT-2000;   2000US-0242332P.
PR   24-OCT-2000;   2000US-0242343P.
PR                  2000US-0243019P.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R, Dang HT, Lowitz KP;
XX
DR   WPI; 2001-355616/37.
DR   N-PSDB; AAS07946.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents.
XX
PS   Claim 45; Page 110-111; 160pp; English.
XX
CC   The sequence represents a human G-protein coupled receptor (GPCR),
CC   hRUP19. The endogenous and non-endogenous, constitutively activated
CC   versions of human G-protein coupled receptors (GPCR), are useful for
CC   direct identification of candidate compounds as receptor agonists,
CC   inverse agonists or partial agonists having applicability as therapeutic
CC   agents for treating diseases related to GPCR, e.g. lung cancer. Non-
CC   endogenous version of human GPCRs are also utilized in research settings
CC   and in vitro and in vivo system, incorporating GPCRs can be utilised to
CC   elucidate and understand the roles these receptors play in the human
CC   condition, both normal and diseased
XX
SQ   Sequence 346 AA;

Query Match     100.0%;  Score 1853;  DB 4;  Length 346;
Best Local Similarity   100.0%;  Pred. No. 5.9e-199;
Matches  346;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60

Qy    61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120

Qy   121  HHAVNTISTRVAAGIVCLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  HHAVNTISTRVAAGIVCLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM   180

Qy   181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240

Qy   241  LYFLWTVPSSACDPSVHGALHITLSFYTMNSMLDPLVYYFSSSPFPKFYNKLKICSLKPK  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  LYFLWTVPSSACDPSVHGALHITLSFYTMNSMLDPLVYYFSSSPFPKFYNKLKICSLKPK  300
```

CC myelofibrosis, psoriasis and cancer including adenocarcinoma, leukaemia,
CC lymphoma; neurological disorders such as epilepsy, ischaemic
CC cerebrovascular disease, Alzheimer's disease, Pick's disease, dementia,
CC Parkinson's disease, ataxias, multiple sclerosis, bacterial and viral
CC meningitis, Creutzfeldt-Jakob disease, schizophrenic disorders, amnesia;
CC cardiovascular disorders such as arteriovenous fistula, atherosclerosis,
CC hypertension, vascular tumours, myocardial infarction, hypertensive heart
CC disease, infective endocarditis, cardiomyopathy, myocarditis;
CC gastrointestinal disorders such as dysphagia, peptic oesophagitis,
CC emesis, anorexia, nausea, peptic ulcer, cholelithiasis, diarrhoea,
CC constipation, acquired immune deficiency syndrome (AIDS), hepatic
CC encephalopathy; autoimmune/inflammatory disorders such as Addison's
CC disease, allergies, spondylitis, amyloidosis, anaemia, asthma, contact
CC dermatitis, Crohn's disease, diabetes mellitus, Goodpasture's syndrome,
CC emphysema, Grave's disease, gout, multiple sclerosis, rheumatoid
CC arthritis, systemic lupus erythematosus, uveitis, viral, bacterial,
CC fungal, parasitic, protozoal and helminthic infections and trauma;
CC metabolic disorders such as diabetes, obesity and osteoporosis; and viral
CC infections such as infection caused by viral agent classified as
CC adenovirus, arenavirus, bunyavirus. Polynucleotides of the invention are
CC useful as probes for assessing toxicity of test compounds. They are also
CC used in gene therapy. The present sequence is human G-protein coupled
CC receptor 3 (GCREC-3) protein
XX
SQ Sequence 346 AA;

Query Match   100.0%; Score 1853; DB 5; Length 346;
Best Local Similarity 100.0%; Pred. No. 5.9e-199;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGPCFHMKTWKPSTVYLFNLAVA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGPCFHMKTWKPSTVYLFNLAVA 60

Qy   61 DFLLMICLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYPKVVHP 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DFLLMICLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYPKVVHP 120

Qy 121 HHAVNTISTRVAAGIVCTLWALVILIGTVTILLENHLCVQETAVSCESFIMESANGWHDIM 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 HHAVNTISTRVAAGIVCTLWALVILIGTVTILLENHLCVQETAVSCESFIMESANGWHDIM 180

Qy 181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240

Qy 241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300

Qy 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346

RESULT 6
AAU11401
ID   AAU11401 standard; protein; 346 AA.
XX
AC   AAU11401;
XX
DT   26-FEB-2002 (first entry)
XX
DE   HM74-like G-protein coupled receptor (GPCR).
XX
KW   HM74-like GPCR; G-protein coupled receptor; antibacterial; fungicide;
KW   protozoacide; analgesic; cytostatic; neuroleptic; nootropic;
KW   anticonvulsant; tranquilliser; viral infection; pain; cancer; anorexia;
KW   bulimia; asthma; central nervous system disease; CNS disease;
KW   cardiovascular disease; hypotension; hypertension; angina pectoris;
KW   myocardial infarction; urinary retention; osteoporosis; ulcer; asthma;
KW   inflammation; allergy; benign prostatic hypertrophy; multiple sclerosis;
KW   psychotic disorder; neurological disorder; dyskinesia;
KW   Huntington's disease; Tourette's syndrome; anxiety; schizophrenia;
KW   manic depression; delirium; dementia; mental retardation.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   Domain   20..37
FT   /label= Transmembrane_domain_1
FT   Domain   53..73
FT   /label= Transmembrane_domain_2
FT   Domain   91..113
FT   /label= Transmembrane_domain_3
FT   Region   101..118
FT   /label= GPCR_region
FT   /note= "G-protein coupled receptor region"
FT   Domain   133..150
FT   /label= Transmembrane_domain_4
FT   Domain   180..197
FT   /label= Transmembrane_domain_5
FT   Domain   223..242
FT   /label= Transmembrane_domain_6
FT   Domain   260..279
FT   /label= Transmembrane_domain_7
XX
PN   WO200177320-A2.
XX
PD   18-OCT-2001.
XX
PF   04-APR-2001; 2001WO-EP003811.
XX
PR   05-APR-2000; 2000US-0194701P.
XX
PA   (FARB) BAYER AG.
XX
PI   Xiao Y;
XX
DR   WPI; 2002-049147/06.
DR   N-PSDB; AAS18501.
XX
PT   Novel isolated polynucleotide, useful for treating infection, pain,
PT   cancer, asthma, hypertension, myocardial infarction, urinary retention,
PT   osteoporosis, encodes the human HM74-like G-protein coupled receptor
PT   polypeptide.
XX
PS   Claim 1; Fig 2; 77pp; English.
XX
CC   The invention describes a novel isolated polynucleotide (I) encoding a
CC   human HM74-like G-protein coupled receptor (GPCR) polypeptide. Reagents
CC   that regulate HM74-like GPCR are useful for modulating the activity of
CC   the protein in a disease selected from bacterial, fungal, protozoan, and
CC   viral infection, pain, cancer, anorexia, bulimia, asthma, central nervous
CC   system (CNS) disease, cardiovascular disease, hypotension, hypertension,
CC   angina pectoris, myocardial infarction, urinary retention, osteoporosis,
CC   ulcer, asthma, inflammation, allergy, benign prostatic hypertrophy,
CC   multiple sclerosis and dyskinesia such as Huntington's disease and
CC   Tourette's syndrome. The composition is also useful for treating
CC   psychotic and neurological disorders such as anxiety, schizophrenia,
CC   manic depression, delirium, dementia and severe mental retardation. (I)
CC   or the HM74-like GPCR polypeptide are also useful for treating the above
CC   mentioned diseases. (I) is useful in a diagnostic assay for detecting
CC   diseases, susceptibility to diseases and abnormalities related to the
CC   presence of mutations in the nucleic acid sequences which encode a GPCR.
CC   The polypeptide is useful to identify test compounds which may act as
CC   agonists or antagonists at the receptor site and which can be regulated
CC   to provide therapeutic effects. The polypeptide is also useful as a bait
CC   protein in a two-hybrid or three-hybrid assay, and to immunise a mammal
CC   for production of polyclonal antibodies. This the amino acid sequence of
CC   human HM74-like GPCR described in the method of the invention
XX
SQ   Sequence 346 AA;

Query Match   100.0%; Score 1853; DB 5; Length 346;
Best Local Similarity 100.0%; Pred. No. 5.9e-199;

Matches 346; Conservative   0;   Mismatches   0;   Indels   0; Gaps   0;

```
Qy      1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60

Qy     61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120

Qy    121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180

Qy    181  FQLEFFMPLGIILFCSFKIWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  FQLEFFMPLGIILFCSFKIWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR   240

Qy    241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK  300

Qy    301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
```

RESULT 7
AAE17077
ID   AAE17077 standard; protein; 346 AA.
XX
AC   AAE17077;
XX
DT   18-APR-2002 (first entry)
XX
DE   Human G-protein coupled receptor (GPCRx14) protein.
XX
KW   Human; G-protein coupled receptor; GPCRx14; cerebroprotective; vomiting;
KW   receptor-mediated disorder; therapy; urinary retention; allergy; obesity;
KW   osteoporosis; angina pectoris; restenosis; atherosclerosis; hypotension;
KW   anorexia; tumour; migraine; acute heart failure; ulcer; antiinflammatory;
KW   stroke; hypertension; neuronal disorder; myocardial infarction psychotic;
KW   depression; mental retardation; neurodegenerative disease; antibacterial;
KW   Alzheimer's disease; dementia; ischaemia; Parkinson's disease; antiviral;
KW   Huntington's disease; anxiety; antifungal; immunosuppressive; cytostatic;
KW   vulnerary; analgesic; anorectic; anabolic; diuretic; cardiant; nootropic;
KW   antiemetic; vasotropic; diabetes; cancer; tranquillizer; neuroleptic.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Domain   17..40
FT   /note= "Transmembrane domain"
FT   Domain   52..70
FT   /note= "Transmembrane domain"
FT   Domain   90..111
FT   /note= "Transmembrane domain"
FT   Domain   132..152
FT   /note= "Transmembrane domain"
FT   Domain   185..203
FT   /note= "Transmembrane domain"
FT   Domain   221..237
FT   /note= "Transmembrane domain"
FT   Domain   258..281
FT   /note= "Transmembrane domain"
XX
PN   WO200198330-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-BE000104.
XX
PR   20-JUN-2000; 2000US-0212913P.

PR   11-JUL-2000; 2000US-0217494P.
PR   26-JAN-2001; 2001EP-00870015.
PR   12-FEB-2001; 2001EP-00870024.
XX
PA   (EURO-) EUROSCREEN SA.
XX
PI   Lannoy V, Brezillon S, Detheux M, Parmentier M, Govarts C;
XX
DR   WPI; 2002-130789/17.
DR   N-PSDB; AAD27497.
XX
PT   New G-protein coupled receptor, useful in the manufacture of medicaments
PT   for treating receptor mediated disorders e.g. acute heart failure and
PT   Alzheimer's disease.
XX
PS   Disclosure; Page 29; 46pp; English.
XX
CC   The present invention relates to a G-protein coupled receptor (GPCR) and
CC   nucleotide encoding it. GPCR are useful in the manufacture of a
CC   medicament for the prevention and/or treatment of receptor-mediated
CC   disorders e.g. viral infections, virus and bacterial diseases, diseases
CC   and disorders involving disturbances of cell migration, diseases or
CC   perturbations of immune system including cancers, development of tumours
CC   and tumour metastasis, inflammatory and neoplastic processes; bacterial
CC   and fungal infections, in wound and bone healing, dysfunction of
CC   regulatory growth functions; pains, diabetes, obesity, anorexia, bulimia,
CC   urinary retention, osteoporosis, angina pectoris, atherosclerosis,
CC   restenosis, diseases involving excessive or reduced proliferation or loss
CC   of smooth muscle cells, aneurysm, stroke, ischaemia, ulcers, allergies,
CC   benign prostatic hypertrophy, migraine, vomiting; blood circulating
CC   affections including acute heart failure, hypotension, hypertension and
CC   myocardial infarction psychotic; neuronal disorders such as anxiety,
CC   schizophrenia, maniac depression, depression, delirium, dementia, severe
CC   mental retardation; degenerative diseases; neurodegenerative diseases
CC   such as Alzheimer's disease, Parkinson's disease; and dyskinesias e.g.
CC   Huntington's disease or Gilles de la Tourette's syndrome and other
CC   related diseases. The present sequence is GPCRx14 protein.
XX
SQ   Sequence 346 AA;

Query Match   100.0%; Score 1853; DB 5; Length 346;
Best Local Similarity   100.0%; Pred. No. 5.9e-199;
Matches 346; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60

Qy     61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120

Qy    121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180

Qy    181  FQLEFFMPLGIILFCSFKIWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  FQLEFFMPLGIILFCSFKIWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR   240

Qy    241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK  300

Qy    301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    301  QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
```

RESULT 8
ABB08596
ID   ABB08596 standard; protein; 346 AA.
XX

```
AC   ABB08596;
XX
DT   01-JUL-2002 (first entry)
XX
DE   Human lipocyte-originated G protein-coupled receptor protein TGR13.
XX
KW   Antiinflammatory; anorectic; obesity; inflammation; gene therapy; human;
KW   G protein-coupled receptor protein TGR13.
XX
OS   Homo sapiens.
XX
PN   WO200202767-A1.
XX
PD   10-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-JP005711.
XX
PR   04-JUL-2000; 2000JP-00206860.
PR   31-JUL-2000; 2000JP-00235274.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Shibata S, Horikoshi K, Taniyama Y, Shintani Y, Miyajima N;
XX
DR   WPI; 2002-164535/21.
DR   N-PSDB; ABA9236.
XX
PT   New human lipocyte-originated G protein-coupled receptor proteins TGR13
PT   and encoding DNAs, for developing drugs to treat obesity and
PT   inflammations, including gene therapy.
XX
PS   Claim 1; Fig 2; 101pp; Japanese.
XX
CC   This invention relates to a human lipocyte-originated G protein-coupled
CC   receptor proteins TGR13, thought to be antiinflammatory and anorectic in
CC   their action. The proteins and encoded DNAs are for use in developing
CC   drugs to treat obesity and inflammation, including gene therapy. The
CC   present sequence represents the human lipocyte-originated G protein-
CC   coupled receptor protein TGR13
XX
SQ   Sequence 346 AA;

Query Match           100.0%;   Score 1853;   DB 5;   Length 346;
Best Local Similarity 100.0%;   Pred. No. 5.9e-199;
Matches 346;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60

Qy     61 DFLIMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 DFLIMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Qy    121 HHAVNTISRVAAGIVCTLWALVIIGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 HHAVNTISRVAAGIVCTLWALVIIGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM 180

Qy    181 FQLRFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FQLRFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240

Qy    241 LYYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPFPKFYNKLKICSLKPK 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 LYYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPFPKFYNKLKICSLKPK 300

Qy    301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
```

```
ID   ABG93786 standard; protein; 346 AA.
XX
AC   ABG93786;
XX
DT   26-NOV-2002 (first entry)
XX
DE   Human G protein-coupled receptor protein, nGPCR-11.
XX
KW   Human; receptor; G protein-coupled receptor; GPCR; nGPCR; beGPCR;
KW   nG protein coupled receptor; communication; serpentine structure;
KW   seven transmembrane receptor; 7TM; mental disorder; diagnosis;
KW   genetic predisposition; brain; immune response; gene therapy;
KW   anxiety disorder; depression; bipolar disorder; schizophrenia;
KW   Huntington's disease; dyskinesia; manic depression; stroke;
KW   Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
KW   tranquiliser.
XX
OS   Homo sapiens.
XX
PN   WO200264789-A1.
XX
PD   22-AUG-2002.
XX
PF   14-FEB-2001; 2001WO-US004641.
XX
PR   14-FEB-2001; 2001WO-US004641.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Lind P, Parodi LA, Vogeli G, Wood LS;
XX
DR   WPI; 2002-674879/72.
DR   N-PSDB; ABS70241.
XX
PT   New nucleic acids and polypeptides of the nG protein-coupled receptor,
PT   useful for treating or diagnosing a mental disorder or a disorder
PT   affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
PT   Parkinson's disease.
XX
PS   Example 1; Page 84; 244pp; English.
XX
CC   The invention discloses an isolated human polypeptide, and encoding
CC   nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
CC   nG protein coupled receptor-14 (nGPCR-14). GPCRs are vital in the
CC   communication between cells and their environment and are characterised
CC   by a serpentine structure that passes through the cell membrane seven
CC   times, hence the reason such receptors are sometimes called seven
CC   transmembrane receptors (7TM). The polynucleotides and polypeptides are
CC   useful for identifying an nGPCR allelic variant that correlates with a
CC   mental disorder, for isolating an antibody that binds to an epitope of
CC   the polypeptide, for identifying a compound that binds the polypeptide or
CC   polynucleotide and/or modulates its biological activity, for screening a
CC   human subject to diagnose a disorder, or a genetic predisposition to a
CC   disorder, affecting the brain or a genetic disposition to the disorder,
CC   for identifying compounds useful for the treatment of a mental disorder
CC   and for identifying a compound useful as a modulator of binding between
CC   nGPCR-14 and a binding partner of nGPCR-14. The polypeptide is also
CC   useful for inducing an immune response in a mammal. The nucleic acid or
CC   polypeptide is particularly useful, using gene therapy, for treating e.g.
CC   anxiety disorders, depression, bipolar disorder, schizophrenia,
CC   Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
CC   disease or Alzheimer's disease. The nucleic acid and polypeptide may also
CC   be used for treating diabetes, inflammation or wounds. The sequences
CC   presented in ABG93747-ABG93793, ABG93795 and ABG93796 are the nGPCR (also
CC   referred to as beGPCRs) proteins
XX
SQ   Sequence 346 AA;

Query Match           100.0%;   Score 1853;   DB 5;   Length 346;
Best Local Similarity 100.0%;   Pred. No. 5.9e-199;
Matches 346;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
```

RESULT 9
ABG93786

```
Db
Qy    1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60

Qy   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120

Qy  121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180

Qy  181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240

Qy  241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNKLKICSLKPK  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNKLKICSLKPK  300

Qy  301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346


RESULT 10
ABP95599
ID   ABP95599 standard; protein; 346 AA.
XX
AC   ABP95599;
XX
DT   06-MAR-2003 (first entry)
XX
DE   Human GPCR polypeptide SEQ ID NO 8.
XX
KW   Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
KW   drug development; gustatory; taste; fragrance; receptor.
XX
OS   Homo sapiens.
XX
PN   WO200216548-A2.
XX
PD   28-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-IB001446.
XX
PR   04-AUG-2000; 2000JP-00237818.
PR   13-FEB-2001; 2001JP-00034434.
XX
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI   Haga T, Takeda S, Mitaku S;
XX
DR   WPI; 2002-304118/34.
DR   N-PSDB; ABZ42873.
XX
PT   Database global search for G protein-coupled receptors, proteins and
PT   encoded genes for studying in vivo signal transduction mechanism and
PT   identifying targets for drug development.
XX
PS   Claim 10; SEQ ID NO 8; 97pp + Sequence Listing; Japanese.
XX
CC   The invention relates to a method for screening G protein-coupled
CC   receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
CC   ABP95942) by extracting open-reading frames containing 6-8 transmembrane
CC   domains with 250-1000 amino acid residues to give a gene homologous with
CC   a known GPCR gene. The receptor proteins and encoded genes are useful for
CC   studying in vivo signal transduction mechanism and identifying targets
CC   for drug development e.g. based on olfactory and gustatory receptors in
CC   form of agonists and antagonists by screening intrinsic and extrinsic
CC   ligands as bitter taste inhibitors, taste enhancers and fragrance
CC   improvers. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
```

```
XX
SQ   Sequence 346 AA;

Query Match             100.0%;  Score 1853;  DB 5;  Length 346;
Best Local Similarity   100.0%;  Pred. No. 5.9e-199;
Matches  346;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA   60

Qy   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120

Qy  121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM  180

Qy  181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240

Qy  241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNKLKICSLKPK  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNKLKICSLKPK  300

Qy  301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346


RESULT 11
AAO14788
ID   AAO14788 standard; protein; 346 AA.
XX
AC   AAO14788;
XX
DT   28-JUN-2002 (first entry)
XX
DE   Human purinergic-like G-protein coupled receptor (AXOR87).
XX
KW   Human; purinergic-like G-protein coupled receptor; AXOR87; immunity;
KW   autoimmunity; inflammation; immunodeficiency; bacterial infection;
KW   fungal infection; viral infection; protozoa infection; cancer; diabetes;
KW   obesity; anorexia; bulimia; asthma; psoriasis; rheumatoid arthritis;
KW   osteoarthritis; psychotic disorder; neurological disorder; vaccine;
KW   chromosome 12q24.
XX
OS   Homo sapiens.
XX
PN   GB2365868-A.
XX
PD   27-FEB-2002.
XX
PF   25-MAY-2001; 2001GB-00012860.
XX
PR   30-MAY-2000; 2000US-00580675.
PR   02-NOV-2000; 2000GB-00026839.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Ignar DM, Elshourbagy N, Gattu M, Shabon U;
XX
DR   WPI; 2002-364852/40.
DR   N-PSDB; AAL42499.
XX
PT   New purinergic-like G-protein coupled receptor AXOR87 polypeptide and
PT   polynucleotide, useful for treating diseases related to autoimmunity,
PT   inflammation, immunodeficiency, or bacterial, fungal, viral and protozoal
PT   infections.
```

```
XX
PS   Claim 2; Page 36; 47pp; English.
XX
CC   The invention comprises the amino acid and coding sequence of the human
CC   purinergic-like G-protein coupled receptor AXOR87. The AXOR87 DNA and
CC   protein sequences of the invention may be used for treating diseases
CC   related to immunity, autoimmunity, inflammation, immunodeficiency, and
CC   infections (i.e. bacterial, fungal, viral, protozoan). The AXOR87 DNA and
CC   protein sequences are particularly useful for treating; cancers,
CC   diabetes, obesity, anorexia, bulimia, asthma, psoriasis, rheumatoid
CC   arthritis, osteoarthritis, as well as psychotic and neurological
CC   disorders. The AXOR87 DNA and protein sequences may also be used as
CC   vaccines. The present amino acid sequence (encoded by a sequence located
CC   on chromosome 12q24) represents the human AXOR87 protein
XX
SQ   Sequence 346 AA;

Query Match      100.0%;  Score 1853;  DB 5;  Length 346;
Best Local Similarity   100.0%;  Pred. No. 5.9e-199;
Matches  346;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 MYNGSSCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MYNGSSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60

Qy   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Qy  121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLENHLCVQBTAVSCESFIMESANGWHDIM 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLENHLCVQBTAVSCESFIMESANGWHDIM 180

Qy  181 FQLEFFMPLGIILFCSFKIVWSLRRRQOLARQARMKKATRFIMWVAIVFITCYLPSVSAR 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 FQLEFFMPLGIILFCSFKIVWSLRRRQOLARQARMKKATRFIMWVAIVFITCYLPSVSAR 240

Qy  241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300

Qy  301 QPGHSKTQRPEBMPISNLGRRSCISVANSFQSQSDGQWDPHIVBWH 346
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  301 QPGHSKTQRPEBMPISNLGRRSCISVANSFQSQSDGQWDPHIVBWH 346

RESULT 12
AAE24354
ID   AAE24354 standard; protein; 346 AA.
XX
AC   AAE24354;
XX
DT   04-OCT-2002 (first entry)
XX
DE   Human G protein coupled receptor (GPCR), 57242 protein.
XX
KW   Human; G protein coupled receptor; GPCR; 57242 protein; overweight;
KW   metabolic disorder; lipogenesis; lipolysis; immunomodulator; heart;
KW   bone disorder; osteoporosis; osteogenesis; bone resorption; cachexia;
KW   hyperlipidaemia; anorexia; haematopoietic disorder; osteopathic;
KW   autoimmune disorder; psoriasis; multiple sclerosis; brain disorder;
KW   degenerative disease; Alzheimer's disease; Pick disease; diabetes;
KW   adipocyte; hyperplastic growth; hypertrophic growth; gene therapy;
KW   obesity; anorectic; receptor.
XX
OS   Homo sapiens.
XX
FH   Key          Location/Qualifiers
FT   Peptide      1..37
FT                /label= Signal_peptide
FT   Domain       1..20
FT                /note= "N-terminal non-transmembrane domain"
FT   Modified-site 3..6
FT   Cleavage-site 9..10
FT                /note= "N-glycosylation site"
FT                /note= "Cleavage site for mitochondrial preseq"
FT   Domain       21..42
FT                /note= "Transmembrane domain"
FT   Domain       32..278
FT                /note= "GPCR domain"
FT   Protein      38..346
FT                /note= "Human mature GPCR protein"
FT   Domain       43..51
FT                /note= "Non-transmembrane domain"
FT   Domain       52..70
FT                /note= "Transmembrane domain"
FT   Region       71..89
FT                /note= "Extracellular loop"
FT   Peptide      77..80
FT                /note= "Nuclear localisation signal"
FT   Domain       90..111
FT                /note= "Transmembrane domain"
FT   Domain       112..130
FT                /note= "Non-transmembrane domain"
FT   Domain       131..152
FT                /note= "Transmembrane domain"
FT   Region       153..184
FT                /note= "Extracellular loop"
FT   Domain       185..201
FT                /note= "Transmembrane domain"
FT   Domain       202..220
FT                /note= "Non-transmembrane domain"
FT   Peptide      204..220
FT                /note= "Nuclear localisation signal"
FT   Modified-site 216..219
FT                /note= "cAMP- and cGMP-dependent protein kinase
FT                phosphorylation site"
FT   Domain       221..245
FT                /note= "Transmembrane domain"
FT   Region       245..258
FT                /note= "Extracellular loop"
FT   Domain       246..255
FT                /note= "Non-transmembrane domain"
FT   Domain       259..280
FT                /note= "Transmembrane domain"
FT   Domain       281..346
FT                /note= "C-terminal cytoplasmic domain"
XX
PN   WO200218579-A2.
XX
PD   07-MAR-2002.
XX
PF   29-AUG-2001; 2001WO-US026882.
XX
PR   29-AUG-2000; 2000US-0228409P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Glucksmann MA;
XX
DR   WPI; 2002-479433/51.
DR   N-PSDB; AAD39181.
XX
PT   Human G protein coupled receptor nucleic acid and polypeptide molecules,
PT   designated 57242, useful for diagnosing, preventing or treating aberrant
PT   lipogenesis or aberrant lipolysis, obesity, diabetes or bone disorders
PT   (e.g. osteoporosis).
XX
PS   Claim 9; Page 112-113; 114pp; English.
XX
CC   The invention relates to G protein coupled receptor (GPCR) family member,
CC   57242 and its corresponding nucleic acid sequence. The 57242 nucleic acid
CC   and polypeptide are useful for diagnosing, preventing or treating a
CC   subject having or at risk of developing a metabolic disorder,
CC   particularly a disorder associated with aberrant lipogenesis or aberrant
CC   lipolysis, obesity or diabetes. The 57242 DNA and protein are also useful
```

```
CC   for treating a subject having bone disorder, where the disorder is
CC   osteoporosis or a disorder associated with aberrant osteogenesis or
CC   aberrant bone resorption. These diseases include obesity, diabetes,
CC   hyperlipidaemia, overweight, anorexia or cachexia. The 57242 DNA and
CC   protein are also useful for treating a subject having haematopoietic
CC   disorders, autoimmune disorders e.g. psoriasis and multiple sclerosis,
CC   brain disorders, degenerative diseases e.g. Alzheimer's disease and Pick
CC   disease and disorders involving heart. The 57242 nucleic acid and
CC   polypeptide are also useful for modulating adipocyte activity such as
CC   hyperplastic growth, hypertrophic growth or lipogenesis. The 57242 DNA is
CC   used in gene therapy. The present sequence is human 57242 protein
XX
SQ   Sequence 346 AA;

Query Match             100.0%;  Score 1853;  DB 5;  Length 346;
Best Local Similarity   100.0%;  Pred. No. 5.9e-199;
Matches  346;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVA  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVA  60

Qy         61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120

Qy        121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCBSFIMESANGWHDIM  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCBSFIMESANGWHDIM  180

Qy        181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240

Qy        241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK  300

Qy        301  QPGHSKTQRPEBMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
               ||||||||||||||||||||||||||||||||||||||||||||||
Db        301  QPGHSKTQRPEBMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346

RESULT 13
ABP81747
ID   ABP81747 standard; protein; 346 AA.
XX
AC   ABP81747;
XX
DT   04-MAR-2003  (first entry)
XX
DE   Human chemokine receptor FKSG80/GPR81 protein SEQ ID NO:668.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer.
XX
OS   Homo sapiens.
XX
PN   WO200261087-A2.
XX
PD   08-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US050107.
XX
```

```
PR   19-DEC-2000; 2000US-0257144P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Burmer GC,  Roush CL,  Brown JP;
XX
DR   WPI; 2003-046718/04.
DR   N-PSDB; ABZ42592.
XX
PT   New isolated antigenic peptides e.g., for G protein-coupled receptors
PT   (GPCR), useful for diagnosing and designing drugs for treating conditions
PT   in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT   autoimmune diseases.
XX
PS   Disclosure; Fig 1; 523pp; English.
XX
CC   The present invention describes antigenic peptides (I) comprising: (a)
CC   any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular G
CC   protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity or
CC   avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC   gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC   antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention
XX
SQ   Sequence 346 AA;

Query Match             100.0%;  Score 1853;  DB 6;  Length 346;
Best Local Similarity   100.0%;  Pred. No. 5.9e-199;
Matches  346;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVA  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1  MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVA  60

Qy         61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61  DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP  120

Qy        121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCBSFIMESANGWHDIM  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121  HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCBSFIMESANGWHDIM  180

Qy        181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181  FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR  240

Qy        241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK  300

Qy        301  QPGHSKTQRPEBMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
               ||||||||||||||||||||||||||||||||||||||||||||||
Db        301  QPGHSKTQRPEBMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH  346
```

RESULT 14
ABP56751
ID   ABP56751 standard; protein; 346 AA.
XX
AC   ABP56751;
XX
DT   31-MAR-2003 (first entry)
XX
DE   Human GAVE3 protein SEQ ID NO:2.
XX
KW   Human; GAVE3; G protein-coupled receptor; GPCR; antiinflammatory;
KW   antiasthmatic; antirheumatic; antiarthritic; inflammatory disorder;
KW   asthma; chronic obstructive pulmonary disease; rheumatoid arthritis;
KW   chromosome 12.
XX
OS   Homo sapiens.
XX
PN   WO2003000846-A2.
XX
PD   03-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019490.
XX
PR   22-JUN-2001; 2001US-00886041.
XX
PA   (AVET) AVENTIS PHARM IKC.
XX
PI   Xia T, Ni D, Eishingdrelo H, Ardati A, Minnich A, Jupp R;
XX
DR   WPI; 2003-184040/18.
DR   N-PSDB; AB522648.
XX
PT   New GAVE3 nucleic acid and polypeptide, for preventing or treating a
PT   disorder associated with aberrant GAVE3 expression or activity, e.g.
PT   inflammatory disease, asthma, chronic obstructive pulmonary disease or
PT   rheumatoid arthritis.
XX
PS   Claim 9; Fig 2; 90pp; English.
XX
CC   The present sequence represents a human G protein-coupled receptor
CC   (GPCR), designated GAVE3. GAVE3 is located on human chromosome 12. GAVE3
CC   has antiinflammatory, antiasthmatic, antirheumatic and antiarthritic
CC   activities. GAVE3 nucleic acids, polypeptides, agonists and antagonists
CC   can be used for preventing or treating a disorder associated with
CC   aberrant GAVE3 expression or activity, e.g. inflammatory disorders such
CC   as asthma, chronic obstructive pulmonary disease or rheumatoid arthritis.
CC   The nucleic acids and polypeptides can also be useful for identifying
CC   modulators, i.e. candidate or test compounds, that bind to GAVE3 proteins
CC   or have a stimulatory or inhibitory effect on GAVE3 expression or
CC   activity. The nucleotide sequences can also be used for chromosomal
CC   mapping, tissue typing or forensic biology. Host cells comprising GAVE3
CC   can be used for producing non-human transgenic animals for studying the
CC   function and/or activity of GAVE3, or for identifying and/or evaluating
CC   modulators of GAVE3 activity
XX
SQ   Sequence 346 AA;

Query Match   100.0%; Score 1853; DB 6; Length 346;
Best Local Similarity   100.0%; Pred. No.5.9e-199;
Matches 346; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60

Qy   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Qy   121 HHAVNTISRVAAGIVCTLWALVIGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 HHAVNTISRVAAGIVCTLWALVIGTVYLLLENHLCVQETAVSCESFIMESANGWHDIM 180

Qy   181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240

Qy   241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSPKFYNKLKICSLKPK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSPKFYNKLKICSLKPK 300

Qy   301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
              |||||||||||||||||||||||||||||||||||||||||||||
Db   301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346

RESULT 15
AAO26511
ID   AAO26511 standard; protein; 346 AA.
XX
AC   AAO26511;
XX
DT   13-FEB-2003 (first entry)
XX
DE   Human G-Protein Coupled Receptor protein.
XX
KW   Anorectic; antidiabetic; antilipaemic; antiasthmatic; antiinflammatory;
KW   antiallergic; antianginal; nephrotropic; hepatotropic; immunosuppressive;
KW   virucide; G-Protein-agonist; G-Protein-antagonist; dyslipidaemia; GPCR;
KW   G-Protein Coupled Receptor; obesity; diabetes; asthma; bronchitis;
KW   allergy; angina; glomerulonephritis; hepatitis; allograft rejection;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO200283736-A2.
XX
PD   24-OCT-2002.
XX
PF   14-FEB-2002; 2002WO-US004397.
XX
PR   14-FEB-2001; 2001US-0269040P.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Elliott SG, Rogers N, Busse LA;
XX
DR   WPI; 2003-073524/07.
DR   N-PSDB; AAL53846.
XX
PT   New GPCR polypeptide and encoding nucleic acid molecule, useful for
PT   diagnosis, treatment and/or prevention of diseases associated with GPCR
PT   polypeptides, such as obesity, diabetes, asthma, allergies, angina and
PT   hepatitis.
XX
PS   Claim 13; Fig 1; 122pp; English.
XX
CC   The invention relates to a novel isolated G-Protein Coupled Receptor
CC   (GPCR) protein. The methods and compositions of the present invention are
CC   useful for diagnosis, treatment, amelioration and/or prevention of
CC   diseases associated with G-Protein Coupled Receptor (GPCR) polypeptides,
CC   such as obesity, diabetes, dyslipidaemia, asthma, bronchitis, allergies,
CC   angina, glomerulonephritis, hepatitis and allograft rejection. This
CC   sequence represents the human GPCR protein of the invention
XX
SQ   Sequence 346 AA;

Query Match   100.0%; Score 1853; DB 6; Length 346;
Best Local Similarity   100.0%; Pred. No. 5.9e-199;
Matches 346; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60

Qy   61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Db   994 TGCCTCCAGAGGAAGATGACAGGTGAGCCAGATA 1027

Search completed: July 3, 2004, 08:42:48
Job time : 3254 secs

```
                    GenCore version 5.1.6

            Copyright (c) 1993 - 2004 Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:     July  3, 2004, 07:48:36 ; Search time 559 Seconds
                                       (without alignments)
                                       8958.010 Million cell updates/sec

Title:      US-10-076-260-1
Perfect score: 1038
Sequence:   1 atgtacaacgggtcgtgctg.........cccacattgttgagtcggcac 1038

Scoring table: IDENTITY_NUC
            Gapop 10.0 , Gapext 1.0

Searched:   3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters:    6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Published Applications NA:*
     1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
     2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
     3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
     4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
     5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
     6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
     7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
     8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
     9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
    10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
    11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
    12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
    13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
    14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
    15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
    16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
    17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
    18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
    19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	
1	1038	100.0	1038	15	US-10-240-842-1	Sequence 1, Appli
2	1038	100.0	1038	15	US-10-076-260-1	Sequence 1, Appli
3	1038	100.0	1038	17	US-10-332-032-2	Sequence 2, Appli
4	1038	100.0	1041	9	US-09-862-274-1	Sequence 1, Appli
5	1038	100.0	1041	10	US-09-886-041-1	Sequence 1, Appli
6	1038	100.0	1041	10	US-09-782-974C-79	Sequence 79, Appli
7	1038	100.0	1041	13	US-10-343-650A-7	Sequence 7, Appli
8	1038	100.0	1041	15	US-10-188-149A-1	Sequence 1, Appli
9	1038	100.0	1041	15	US-10-079-384-17	Sequence 17, Appl
10	1038	100.0	1041	15	US-10-321-807-23	Sequence 23, Appl
11	1038	100.0	1050	13	US-10-210-172-177	Sequence 177, App
12	1038	100.0	1050	15	US-10-044-643-1	Sequence 1, Appli
13	1038	100.0	1050	15	US-10-044-643-3	Sequence 3, Appli
14	1038	100.0	1057	13	US-10-210-172-183	Sequence 183, App
15	1038	100.0	1057	13	US-10-210-172-191	Sequence 191, App
16	1038	100.0	1060	13	US-10-210-172-187	Sequence 187, App
17	1038	100.0	1083	15	US-10-278-141-11	Sequence 11, Appl
18	1038	100.0	1083	16	US-10-296-081-11	Sequence 11, Appl
19	1038	100.0	1372	15	US-10-225-567A-667	Sequence 667, App
20	1038	100.0	1730	15	US-10-240-842-3	Sequence 3, Appli
21	1038	100.0	2331	15	US-10-188-149A-3	Sequence 3, Appli
22	1038	100.0	2345	15	US-10-201-481-6	Sequence 6, Appli
23	1038	100.0	3612	15	US-10-201-481-5	Sequence 5, Appli
24	1038	100.0	27893	15	US-10-017-161-757	Sequence 757, App
25	1036.4	99.8	1041	9	US-09-942-374-3	Sequence 3, Appli
26	1036.4	99.8	1041	17	US-10-665-956-3	Sequence 3, Appli
27	1036.4	99.8	1104	13	US-10-210-172-179	Sequence 179, App
28	1036.4	99.8	1194	9	US-09-942-374-1	Sequence 1, Appli
29	1036.4	99.8	1194	17	US-10-665-956-1	Sequence 1, Appli
30	1034.8	99.7	1041	15	US-10-094-417-7	Sequence 7, Appli
31	1034.8	99.7	1104	13	US-10-210-172-181	Sequence 181, App
32	1034.8	99.7	1104	15	US-10-044-643-4	Sequence 4, Appli
33	1015.8	97.7	3978	13	US-10-017-161-767	Sequence 767, App
34	1014	97.7	2580	15	US-10-092-135-1	Sequence 1, Appli
35	942.4	90.8	961	13	US-10-210-172-189	Sequence 189, App
36	942	90.8	961	13	US-10-210-172-185	Sequence 185, App
37	942	90.8	961	13	US-10-210-172-193	Sequence 193, App
38	888	85.5	888	10	US-09-782-974C-11	Sequence 11, Appl
39	724.4	69.8	1668	15	US-10-076-260-5	Sequence 5, Appli
40	723.4	69.7	742	13	US-10-210-172-195	Sequence 195, App
41	722.8	69.6	3251	15	US-10-076-260-3	Sequence 3, Appli
42	627.4	60.4	646	13	US-10-210-172-197	Sequence 197, App
43	627.4	60.4	646	13	US-10-210-172-203	Sequence 203, App
44	372.4	35.9	1174	10	US-09-930-334-1	Sequence 1, Appli
45	372.4	35.9	1174	13	US-10-619-141-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-10-240-842-1
; Sequence 1, Application US/10240842
; Publication No. US20030109673A1
; GENERAL INFORMATION:
;  APPLICANT: Xiao, Yonghong
;  TITLE OF INVENTION: REGULATION OF HUMAN HM74-LIKE G PROTEIN-COUPLED RECEPTOR
;  FILE REFERENCE: 4974.00883
;  CURRENT APPLICATION NUMBER: US/10/240,842
;  CURRENT FILING DATE: 2002-10-04
;  PRIOR APPLICATION NUMBER: 60/194,701
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 1
;  LENGTH: 1038
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-240-842-1

Query Match             100.0%;  Score 1038;  DB 15;  Length 1038;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1038;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 ATGTACAACGGGTCGTGCTGCTGCCGCATGCGAGGGGGACACCATCTCCCAGGTGATGCGCCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGTACAACGGGTCGTGCTGCTGCCGCATGCGAGGGGGACACCATCTCCCAGGTGATGCGCCG 60

Qy    61 CTGCTCATTGTGCGCCTTTGTGCTGGGCGGCACTAGGCAATGGGGTCGCCCTGTGCCCTGTGGTTTC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGCTCATTGTGCGCCTTTGTGCTGGGCGGCACTAGGCAATGGGGTCGCCCTGTGCCCTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTACCTTTCAATTGGCCGTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTACCTTTCAATTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCGCCGCTGCCTTTTCGGACAGAGACTATTACCTCAGACGACGTAGACAC 240
```

```
;  PRIOR FILING DATE: 2001-02-14
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 1
;  LENGTH: 1038
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)..(1038)
US-10-076-260-1

Query Match      100.0%;   Score 1038;   DB 15;   Length 1038;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    1  ATGTACAACGGGTGTGCTGCCGCCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGTACAACGGGTGTGCTGCCGCCATCGAGGGGACACCATCTCCAGGTGATGCCGCCG   60

Qy   61  CTGCTCATTGTGGGCCTTTGTCTGGGCGCACTAGGCAATGGGTCGCCCTGTGTGGTTTC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  CTGCTCATTGTGGGCCTTTGTCTGGGCGCACTAGGCAATGGGTCGCCCTGTGTGGTTTC  120

Qy  121  TGCTTCCACATGAAGACCTGGAGCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT     180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  TGCTTCCACATGAAGACCTGGAGCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT     180

Qy  181  GATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GATTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240

Qy  241  TGGGCTTTTGGGGACATTCCCTGCCCGAGTGGGGCTGTCGCGGACAGGTGGGGCC      300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TGGGCTTTTGGGGACATTCCCTGCCCGAGTGGGGCTGTCGCGGACAGGTGGGGCC      300

Qy  301  GGGAGCATCGTGTTCCTTACGGTGTGGTCTGCGGACAGGTATTTCAAAGTGGTCCACCCTG  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GGGAGCATCGTGTTCCTTACGGTGTGGTCTGCGGACAGGTATTTCAAAGTGGTCCACCCC   360

Qy  361  CACCAGCGGTGAACACTATCTCCACCCGGTGGCGGCTGGCATCGTCTGCACCCTGTGG   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CACCAGCGGTGAACACTATCTCCACCCGGTGGCGGCTGGCATCGTCTGCACCCTGTGG   420

Qy  421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGCTGGAGAACCATCTCGCCTGTGG   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGCTGGAGAACCATCTCGCCTGTGG   480

Qy  481  ACGGCCGTCTCCTGTGAGAGTCATCATGAGTCGGCCAATGCTGGCATCATG        540
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ACGGCCGTCTCCTGTGAGAGTCATCATGAGTCGGCCAATGCTGGCATCATG        540

Qy  541  TTCCAGCTGGAGTTCTTTATGCCCCTGCGGATCCCTCTTATTTTGCTCCCTCAAGATTGTT  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  TTCCAGCTGGAGTTCTTTATGCCCCTGCGGATCCCTCTTATTTTGCTCCCTCAAGATTGTT  600

Qy  601  TGGAGCCTGAGGCGGAGGCAGCAGCTGGCAGCAGGCGACCCG                 660
         ||||||||||||||||||||||||||||||||||||||||||
Db  601  TGGAGCCTGAGGCGGAGGCAGCAGCTGGCAGCAGGCGACCCG                 660

Qy  661  TTCATCATGGTGGTGGCAATTGTTCATCACAGCTACCTGCCCAGCGTGTCGCTAGA    720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  TTCATCATGGTGGTGGCAATTGTTCATCACAGCTACCTGCCCAGCGTGTCGCTAGA    720

Qy  721  CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGTCCCGTCCATGGGGCCTG     780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGTCCCGTCCATGGGGCCTG     780

Qy  781  CACATAACCCTCAGCTTCACCTACACATGAACACAGCTGGATCCCCGGTGTATTATTT   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  CACATAACCCTCAGCTTCACCTACACATGAACACAGCTGGATCCCCGGTGTATTATTT   840

Qy  841  TCAAGCCCCTCCTTTCCCAAATTCTACAACAGCTCAAAATTCTGAAACCCAAG       900
```

Db 841 TCAAGCCCCTCCTTTCCACAACAGCTTCACAACAGAGCTGGAAGAGATGCCGAATTTCGACGTCTGAAACCCAAG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 901 CAGCCAGGAGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGACCTCGGTCGC 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 CAGCCAGGAGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGACCTCGGTCGC 960
Qy 961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
Qy 1021 CACATGTTGAGTGGCAC 1038
        |||||||||||||||||
Db 1021 CACATGTTGAGTGGCAC 1038

RESULT 3
US-10-332-032-2
; Sequence 2, Application US/10332032
; Publication No. US2004008694DA1
; GENERAL INFORMATION:
;   APPLICANT: Takeda Chemical Industries, Ltd.
;   TITLE OF INVENTION: Novel G Protein-Coupled Receptor Protein and its DNA
;   FILE REFERENCE: P2001-154PCT
;   CURRENT APPLICATION NUMBER: US/10/332,032
;   CURRENT FILING DATE: 2003-01-03
;   PRIOR APPLICATION NUMBER: JP 2000-206860
;   PRIOR FILING DATE: 2000-07-04
;   PRIOR APPLICATION NUMBER: JP 2000-235274
;   PRIOR FILING DATE: 2000-07-31
;   NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 2
;   LENGTH: 1038
;   TYPE: DNA
;   ORGANISM: Human
US-10-332-032-2

Query Match   100.0%; Score 1038; DB 17; Length 1038;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1038; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCGCCG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCGCCG 60
Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
Qy 121 TGCTTCCACATGAGAAGCCCAGCACTGTTTACCTTTCAATTTGGCCGTGCT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TGCTTCCACATGAGAAGCCCAGCACTGTTTACCTTTCAATTTGGCCGTGCT 180
Qy 181 GATTTCTCCTATGATCGCGCTCCTGCTTTTCGGACAGACTATTACCTCAGACGAGACAC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GATTTCTCCTATGATCGCGCTCCTGCTTTTCGGACAGACTATTACCTCAGACGAGACAC 240
Qy 241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGGC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
Qy 301 GGGAGCAATCGTGTTCCTTACGGTGGTGCTGCGGACAGGTATTCAAAGTGGTCCACCCC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GGGAGCAATCGTGTTCCTTACGGTGGTGCTGCGGACAGGTATTCAAAGTGGTCCACCCC 360
Qy 361 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGCACCCTGTGG 420
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGCACCCTGTGG 420
Qy 421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGTCCAAGAG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGTCCAAGAG 480

Db 481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCATGACTACATCATG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCATGACTACATCATG 540
Db 541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTCAAGATTGTT 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTCAAGATTGTT 630
Db 601 TGGAGCCTGAGGCGGAGGCAGCTGGCCAGACAGGCTCGGATGAAGAGGGCGACCCGG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 601 TGGAGCCTGAGGCGGAGGCAGCTGGCCAGACAGGCTCGGATGAAGAGGGCGACCCGG 660
Db 661 TTCATCATGGTGGTGCCAATTGGTTGTTCATCACATGCTCACCTGCCCCAGCGTGTCTGCTAGA 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 661 TTCATCATGGTGGTGCCAATTGGTTGTTCATCACATGCTCACCTGCCCCAGCGTGTCTGCTAGA 720
Db 721 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTCGGATCCCTCGTCCTGTGTATTATTTT 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 721 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTCGGATCCCTCGTCCTGTGTATTATTTT 780
Db 781 CACATAACCCTCAGCTTCACCTACACAGACAGCATGCTGGATCCCCGTGGTGTATTATTTT 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 781 CACATAACCCTCAGCTTCACCTACACAGACAGCATGCTGGATCCCCGTGGTGTATTATTTT 840
Db 841 TCAAGCCCCTCCTTTCCACAACAGCTTCACAACAGCTGGAAGAGATGCCAATTTCGAGTCTGAAACCCAAG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 841 TCAAGCCCCTCCTTTCCAAATTCTACAACAGCTGGAAGAGATGCCAATTTCGAGTCTGAAACCCAAG 900
Db 901 CAGCCAGGAGACACTCAAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAGTCTGAAACCCAAG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 901 CAGCCAGGACACTCAAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
Db 961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
Db 1021 CACATGTTGAGTGGCAC 1038
        |||||||||||||||||
Qy 1021 CACATGTTGAGTGGCAC 1038

RESULT 4
US-09-862-274-1
; Sequence 1, Application US/09862274
; Patent No. US20020052022A1
; GENERAL INFORMATION:
;   APPLICANT: ELSHOURBAGY, NABIL
;   APPLICANT: GATTU, MAHANANDEESHWAR
;   APPLICANT: SHABON, USMAN
;   APPLICANT: IGNAR, DIANE MICHELE
;   TITLE OF INVENTION: MOLECULAR CLONING OF A CHEMOKINE LIKE
;   TITLE OF INVENTION: 7TMR (AXOR87)
;   FILE REFERENCE: GP-70703-1
;   CURRENT APPLICATION NUMBER: US/09/862,274
;   CURRENT FILING DATE: 2001-05-22
;   PRIOR APPLICATION NUMBER: US 09/580,675
;   PRIOR FILING DATE: 2000-05-30
;   PRIOR APPLICATION NUMBER: GB 0026839.1
;   PRIOR FILING DATE: 2000-11-02
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1041
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-862-274-1

Query Match   100.0%; Score 1038; DB 9; Length 1041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1038; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCGCCG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
;   Sequence 1, Application US/09886041
;   Publication No. US20030059869A1
;   GENERAL INFORMATION:
;   APPLICANT: XIA, TAI-HE
;   APPLICANT: NI, DONGHUI
;   APPLICANT: EISHINGDRELO, HAIFENG
;   APPLICANT: ARDATI, ALI
;   APPLICANT: MINNICH, ANNE
;   APPLICANT: JUPP, RAY
;   TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR
;   FILE REFERENCE: 41491
;   CURRENT APPLICATION NUMBER: US/09/886,041
;   CURRENT FILING DATE: 2001-06-22
;   NUMBER OF SEQ ID NOS: 12
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 1
;   LENGTH: 1041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-886-041-1

Query Match      100.0%;   Score 1038;   DB 10;   Length 1041;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1038;  Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 ATGTACAACGGGTCGTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG    60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGTACAACGGGTCGTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG    60

Qy    61 CTGCTCATTGTGGCCTTTGTGCTGGGGCGACTAGGCAATGGGGGTCGCCCTGTGTGGTTC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGCTCATTGTGGCCTTTGTGCTGGGGCGACTAGGCAATGGGGGTCGCCCTGTGTGGTTC  120

Qy   121 TGCTTCCACATGAAGAACCTGGACACTGTTTACCTTTTCAATTTGGCCGGTGGCT       180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGAACCTGGACACTGTTTACCTTTTCAATTTGGCCGGTGGCT       180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC   240

Qy   241 TGGGCTTTTGGGGACATTCCTGCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC    300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCTGCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC    300

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGCTCGGACAGTATTCAAAGTGGTCCACCCC       360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTTCCTTACGGTGGCTCGGACAGTATTCAAAGTGGTCCACCCC       360

Qy   361 CACCACCGCGTGAACACTATCTCCACCCGGGTGGCGGCTGGACATCGTCTGCACCCTGTGG  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACCGCGTGAACACTATCTCCACCCGGGTGGCGGCTGGACATCGTCTGCACCCTGTGG  420

Qy   421 GCCCTGGTCATCCTGGAACAGTGTATCTTTTGCTGGAGAACATCTTTTGCTGGAGAACATCATG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGAACAGTGTATCTTTTGCTGGAGAACATCTTTTGCTGGAGAACATCATG  480

Qy   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG           540
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG           540

Qy   541 TTCCAGTCGGAGTTCTTTATGCCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGTCGGAGTTCTTTATGCCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600

Qy   601 TGGAGCCTGAGGCGAGGCAGCAGCTGGCAAGTGTTCATCACATG              660
         |||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGAGGCAGCAGCTGGCAAGTGTTCATCACATG              660

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACTACCTGCCCCAGCGTGTCGCCTAGA    720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACTACCTGCCCCAGCGTGTCGCCTAGA    720
```

```
Db       ATGTACAACGGGTCGTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG      60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       CTGCTCATTGTGGCCTTTGTGCTGGGGCGACTAGGCAATGGGGGTCGCCCTGTGTGGTTTC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       CTGCTCATTGTGGGCCTTTGTGCTGGGGCGACTAGGCAATGGGGGTCGCCCTGTGTGGTTTC  120

Qy       TGCTTCCACATGAAGAACCTGGACACTGTTTACCTTTTCAATTTGGCCGGTGGCT       180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       TGCTTCCACATGAAGAAGAACCTGGACACTGTTTACCTTTTCAATTTGGCCGGTGGCT       180

Qy       GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC   240

Qy       TGGGCTTTTGGGGACATTCCCTGCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       TGGGCTTTTGGGGACATTCCCTGCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC   300

Qy       GGGAGCATCGTGTTCCTTACGGTGGCTCGGACAGTATTCAAAGTGGTCCACCCC       360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       GGGAGCATCGTGTTCCTTACGGTGGCTCGGACAGTATTCAAAGTGGTCCACCCC       360

Qy       CACCACCGCGTGAACACTATCTCCACCCGGGTGGCGGCTGGACATCGTCTGCACCCTGTGG  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       CACCACCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGACATCGTCTGCACCCTGTGG  420

Qy       GCCCTGGTCATCCTTTTCCAAATTCACAACCAAGG                         480
         ||||||||||||||||||||||||||||||||||
Db       GCCCTGGTCATCCTTTTCCAAATTCACAACCAAGG                         480

Qy       ACGGCCGTCTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG             540
         |||||||||||||||||||||||||||||||||||||||||||||||
Db       ACGGCCGTCTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG             540

Qy       TTCCAGTCGTCATCAGTTGGCAAATAGTTTGCAAATGGGCAATGGGATCCC         600
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       TTCCAGTCGTCATCAGTTGGCAAATAGTTTGCAAATGGGCAATGGGATCCC         600

Qy       TGGAGCCTGAGGCGAGGCAGCAGCTGGCAAGTGTTCATCACATG               660
         ||||||||||||||||||||||||||||||||||||||||||||
Db       TGGAGCCTGAGGCGAGGCAGCAGCTGGCAAATTGTTCATCACATG              660

Qy       CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGATCCCTGTGTCCATGGGGCCCTG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGATCCCTGTGTCCATGGGGCCCTG  720

Qy       CACATAACCCTCAGCTTCACCTACAGAACATGAACAGACAGCATGGTGTATTATTTT     840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       CACATAACCCTCAGCTTCACCTACAGAACATGAACAGACAGCATGGTGTATTATTTT     840

Qy       TCAAGCCCCTCCTTTCCAAATTCACAACAAGTCGCAGTCTGAAACCCAAG          900
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       TCAAGCCCCTCCTTTCCAAATTCACAACAAGTCGCAGTCTGAAACCCAAG          900

Qy       CAGGCAGGACACTCAAAAACAAAAGGCCGGAAGAGATGCCAATTGGAACCTCGGTCGC     960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       CAGGCAGGACACTCAAAAACAAAAGGCCGGAAGAGATGCCAATTGGAACCTCGGTCGC     960

Qy       AGGAGTTGCATCAGTTGGCAAATAGTTTGCAAATGGGCAATGGGATCCC          1020
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db       AGGAGTTGCATCAGTTGGCAAATAGTTTGCAAATGGGCAATGGGATCCC          1020

Qy       CACATTGTTGAGTGGCAC                                         1038
         |||||||||||||||||
Db       CACATTGTTGAGTGGCAC                                         1038
```

```
RESULT 5
US-09-886-041-1
```

```
Qy      1 ATGTACAACGGGTCGTGCTGCCGCCATCCGAGGGGACACCACCATCTCCCAGTGATGCCGCCG  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        ...

Qy     61 CTGCTCATTGTGGCCTTTGTGCTGTGGGGCCACTAGGCAATGGCAATGGGGTCGCCCTGTGTGGTTTC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCTCATTGTGGCCTTTGTGCTGTGGGGCCACTAGGCAATGGCAATGGGGTCGCCCTGTGTGGTTTC 120

Qy    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180

Qy    181 GATTTCCTCCTTATGATCTGCGCCTTTCGACAGACTATTACCTCAGACGTAGACAC 240
          |||||||||||||||||||||||||||||||||||||||||
Db    181 GATTTCCTCCTTATGATCTGCGCCTTTCGGACACAGACTATTACCTCAGACGTAGACAC 240

Qy    241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGAGGCC 300
          |||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTTTGGGGACACTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGAGGCC 300

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC 360
          |||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC 360

Qy    361 CACCACGCGGTGAACACTATCTCCACCCGGTGGCCGGCTGGCCATCGTCTCGCACCCTGTGG 420
          |||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGCGGTGAACACTATCTCCACCCGGTGGCCGGCTGGCCATCGTCTCGCACCCTGTGG 420

Qy    421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCGGCTGCGCAAGAG 480
          |||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGGAGAACCATCTCGGCTGCGCAAGAG 480

Qy    481 ACGGCCGTCTCCTGTGAGAGCTT GAGAGCTTGCGGCCAATGGCTGGCTGGCATGACATCATG 540
          |||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCCTGTGAGAGCTTGAGAGCTTGCGGCCAATGGCTGGCTGGCATGACATCATG 540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT 600
          |||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT 600

Qy    601 TGGAGCCTGGAGGCCGGAGCGAGCTGGCCAGACAGGCTGGATGAAGAGAGGCGACCCGG 660
          |||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGGAGGCCGGAGCGAGCTGGCCAGACAGGCTGGATGAAGAGAGGCGACCCGG 660

Qy    661 TTCATCATGGTGTGGCCAATTGTGTTCATCACATGCTACCTGCCCAGCCGTGTCTCGCTAGA 720
          |||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGTGGCCAATTGTGTTCATCACATGCTACCTGCCCAGCCGTGTCTCGCTAGA 720

Qy    721 CTCTATTTCCTCTGGACGGTGCCCTGAGTGCCCTGTCCATGGGGGCCTG 780
          |||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGGACGGTGCCCTGAGTGCCCTGTCCATGGGGGCCTG 780

Qy    781 CACATAACCCTCAGCTTCACCTACACAGACAGCATGCTGGATCCCTGGTGTTATTATTTT 840
          |||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCAGCTTCACCTACACAGACAGCATGCTGGATCCCTGGTGTTATTATTTT 840

Qy    841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
          |||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
          |||||||||||||||||||||||||||||||||||||||||||
Db    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy    961 AGGAGTTGCATCAGTGTGGCCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
          |||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGAGTTGCATCAGTGTGGCCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038
```

RESULT 6
US-09-782-974C-79
; Sequence 79, Application US/09782974C
; Publication No. US2003008253 4A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534Alel G Protein Coupled Receptor
; FILE REFERENCE: 411USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-974C-79

Query Match   100.0%;   Score 1038;   DB 10;   Length 1041;
Best Local Similarity   100.0%;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ATGTACAACGGGTCGTGCTGCCGCCATCCGAGGGGACACCACCATCTCCCAGTGATGCCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

; Sequence 7, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
;   APPLICANT: HAGA, TATSUYA
;   TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
;   FILE REFERENCE: 31671-186347
;   CURRENT APPLICATION NUMBER: US/10/343,650A
;   CURRENT FILING DATE: 2003-07-21
;   PRIOR APPLICATION NUMBER: JP 2000/237818
;   PRIOR FILING DATE: 2000-08-04
;   PRIOR APPLICATION NUMBER: JP 2001/34434
;   PRIOR FILING DATE: 2001-02-13
;   NUMBER OF SEQ ID NOS: 694
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 1041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1041)
US-10-343-650A-7

Query Match   100.0%; Score 1038; DB 13; Length 1041;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1038; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGTGATGCCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGTGATGCCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCCACTAGGCAATGGGCGTCGCCCTGTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCCACTAGGCAATGGGCGTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACCACTGTTTACCTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACCACTGTTTACCTTTCAATTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTATGATCTGCCTGCCTTTTCGGACACGACTATTACCTCAGACTAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTATGATCTGCCTGCCTTTTCGGACACGACTATTACCTCAGACTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCCGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCCGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGTCGGCTGCGGACAGGTATTTCAAAGTGTCCACCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTTCCTTACGGTGGTCGGCTGCGGACAGGTATTTCAAAGTGTCCACCCC 360

Qy   361 CACCACGCGCGTGAACACTATCTCACCCGGGTGGCGGCTGGCATCGTCTGCGCACCCTGTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGCGTGAACACTATCTCACCCGGGTGGCGGCTGGCATCGTCTGCGCACCCTGTGG 420

Qy   421 GCCCGGCCGTCTCCGTGTGAGAGCTTCAATCATGGACTGGCAATGGCTGGCATGACATCATG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCGGCCGTCTCCGTGTGAGAGCTTCAATCATGGACTGGCAATGGCTGGCATGACATCATG 480

Qy   481 ACGGGCCGTCTCCTGTGAGAGCTTCATCATGGACTGGCATGTTGCTTCCTTCAAGATTGTT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGGCCGTCTCCTGTGAGAGCTTCATCATGGACTGGCATGTTGCTTCCTTCAAGATTGTT 540

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTGCTTCCTTCAAGATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTGCTTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCAGCAGCCAGCACGTCGCCAGAGAGGCCGACCCGG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCAGCAGCCAGCACGTCGCCAGAGAGGCCGACCCGG 660

Qy   661 TTCATCATGGTGGTGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   661 TTCATCATGGTGGTGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 CTCTATTTCCTCTGGACGGTGCCCTGGACGTGCCTGGGATCCCTCGTCCTCGTCCATGGGGCCCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTCTATTTCCTCTGGACGGTGCCCTGGACGTGCCTGGGATCCCTCGTCCTCGTCCATGGGGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACCTACATGAACACAGCATGCTGGGATCCCTCGTGTTATTATTTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CACATAACCCTCAGCTTCACCTACATGAACACAGCATGCTGGGATCCCTCGTGTTATTATTTT 840

Qy   841 TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 8
US-10-188-149A-1
; Sequence 1, Application US/10188149A
; Publication No. US20030077734A1
; GENERAL INFORMATION:
;   APPLICANT: YE, Jane et al.
;   TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
;   TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
;   TITLE OF INVENTION: PROTEINS, AND USES THEREOF
;   FILE REFERENCE: CL000782-CON
;   CURRENT APPLICATION NUMBER: US/10/188,149A
;   CURRENT FILING DATE: 2002-12-10
;   PRIOR APPLICATION NUMBER: US 09/666,535
;   PRIOR FILING DATE: 2000-09-20
;   PRIOR APPLICATION NUMBER: US 60/230,459
;   PRIOR FILING DATE: 2000-09-06
;   PRIOR APPLICATION NUMBER: US 60/192,419
;   PRIOR FILING DATE: 2000-03-27
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1041
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-188-149A-1

Query Match   100.0%; Score 1038; DB 15; Length 1041;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1038; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCCACTAGGCAATGGGCGTCGCCCTGTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCCACTAGGCAATGGGCGTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACCACTGTTTACCTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACCACTGTTTACCTTTCAATTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGAGACAC 240

```
;       TYPE: DNA
;       ORGANISM: Homo sapiens
;       FEATURE:
;       NAME/KEY: CDS
;       LOCATION: (1)..(1041)
;       OTHER INFORMATION:
US-10-079-384-17

Query Match             100.0%;  Score 1038;  DB 15;  Length 1041;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1038;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy      1 ATGTACAACGGGTCGTGCTGCCGAGTGGGCGTCTTCACGTTGGCCATGAACCAGGTGACACTCCCAGGTGATGCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTACAACGGGTCGTGCTGCCGAGTGGGCGTCTTCACGTTGGCCATGAACCAGGTGACACTCCCAGGTGATGCGCCG 60

Qy     61 CTGCTCATTGTGGCCTTGTCTGCGGCGCACTAGGCAATGGGGTCGCCGTGTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCTCATTGTGGCCTTGTCTGCGGCGCACTAGGCAATGGGGTCGCCGTGTTC 120

Qy    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTACCTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTACCTTTCAATTGGCCGTGGCT 180

Qy    181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy    241 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTTCACGTTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTTCACGTTGGCCATGAACAGGGCC 300

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGCGGGCTGGCGGCCCTG 360
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGCGGGCTGGCGGCCCTG 360

Qy    361 CACCACGGCGGTGAACACTATCTCCACCGGGTGCGGCTGGCATCGTCTGCACCCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGGCGGTGAACACTATCTCCACCGGGTGCGGCTGGCATCGTCTGCACCCC 420

Qy    421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGAGAACCATCTCTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTGAGAACCATCTCTGCGTGCAAGAG 480

Qy    481 ACGGCCGTCTCCTGGACGTGCCCTGCGATCCCGAGTGCTGGCCAATGGCTGGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCCTGGACGTGCCCTGCGATCCCGAGTGCTGGCCAATGGCTGGCATG 540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTCTTATTGTT 600
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTCTTATTGTT 600

Qy    601 TGGAGCCTGAGGCGGAGGCGCAGCAGCTGCCAGCAGAGGCGGCACCCG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGGCGGAGGCGCAGCAGCTGCCAGCAGAGGCGGCACCCG 660

Qy    661 TTCATCATGGTGGTGCCAATTGTGTTCATCACATGCTACCTGCTACCTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGGTGCCAATTGTGTTCATCACATGCTACCTGCTACCTG 720

Qy    721 CTCTATTTCCTCTGGACGTGCCCTGGACGTGCCCTGCGATCCCTGTCG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGGACGTGCCCTGGACGTGCCCTGCGATCCCTGTCG 780

Qy    781 CACATAACCCTGCAAATTCACAACAAGCTCAAAATCTGACAGATCGGAT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCCAGCTTCACCTACAACAAGCTCAAAATCTGACAGATC 840

Qy    841 TCAAGCCCCTCCTTCCCAAATTCACAACAAGCTCAAAATCTCGAAACCCAAG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGCCCCTCCTTCCCAAATTCACAACAAGCTCAAAATCTCGAAACCCAAG 900

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
```

RESULT 9
US-10-079-384-17
;   Sequence 17, Application US/10079384
;   Publication No. US20030108986A1
;   GENERAL INFORMATION:
;   APPLICANT: Communi, Didier
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS COMPRISING G-PROTEIN COUPLED RECEPTORS
;   FILE REFERENCE: 9409/2132
;   CURRENT APPLICATION NUMBER: US/10/079,384
;   CURRENT FILING DATE: 2002-02-20
;   PRIOR APPLICATION NUMBER: US 09/885,453
;   PRIOR FILING DATE: 2001-06-20
;   NUMBER OF SEQ ID NOS: 50
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 17
;   LENGTH: 1041

```
Qy    241 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTTCTTCACGTTGGCCATGAACAGGGCC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTT-TGGGACATTCCCTGCCGAGTGGGGCTTCTTCACGTTGGCCATGAACAGGGCC 300

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGCGGGCTGGCGGCCCTG 360
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGCGGGCTGGCGGCCCTG 360

Qy    361 CACCACGGCGGTGAACACTATCTCCACCGGGTGCGGCATCGTCTGCACCCCGTGG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGGCGGTGAACACTATCTCCACCGGGTGCGGCATCGTCTGCACCCCGTGG 420

Qy    421 GCCCTGGTCATCCTTGCTGTGGAGAACCATCTTTGCTGTGGAGAACCATCTCTGCGTGCAAGAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCTGGTCATCCTTGCTGTGGAGAACCATCTTTGCTGTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy    481 ACGGCCGTCTCCTGAGAGCTTCATCATGGAGTCGGCCAATGGCATCATG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCCTGAGAGCTTCATCATGGAGTCGGCCAATGGCATCATG 540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTGTATTATTTT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTGTATTATTTT 600

Qy    601 TGGAGCCTGAGGCGGAGGCGCAGCAGCTGCCAGCAGAGAGGCGGCACCCG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGGCGGAGGCGCAGCAGCTGCCAGCAGAGAGGCGGCACCCG 660

Qy    661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCTACCTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCTACCTG 720

Qy    721 CTCTATTTCCTCTGAGTGCCCTGAGTGCCCTGCGATCCCTGTCG 780
          ||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGAGTGCCCTGAGTGCCCTGCGATCCCTGTCG 780

Qy    781 CACATAACCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGAAACCCAAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGAAACCCAAG 840

Qy    841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTCGGTCGC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTCGGTCGC 900

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGATCCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038
```

```
Db   901  CAGCCAGGAGCACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTGCGC  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   241  TGGGCTTTTGGGGACATTCCCTGCCGGAGTGGGGCTCTCACGTTGGCCATGAACAGGGCC   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TGGGCTTTTGGGGACATTCCCTGCCGGAGTGGGGCTCTCACGTTGGCCATGAACAGGGCC   300

Qy   301  GGGAGCATCGTGTTCCTTACGGTGGTCGCTGCGGACAGGTATTCAAAGTGGTCCACCCC   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GGGAGCATCGTGTTCCTTACGGTGGTCGCTGCGGACAGGTATTCAAAGTGGTCCACCCC   360

Qy   361  CACCAGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CACCAGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGG   420

Qy   421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG   480

Qy   481  ACGGCCGTCTCCGTGGAGAGCTTCATCATGGCTGGCATGACACATCATG   540
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ACGGCCGTCTCCGTGGAGAGCTTCATCATGGCTGGCATGACACATCATG   540

Qy   541  TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT   600

Qy   601  TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG   660

Qy   661  TTCATCATGGTGGTCGGCAATTGTGTTCATCACATGTTACCTGCCCAGCGTGTCTGCTAGA   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TTCATCATGGTGGTCGGCAATTGTGTTCATCACATGTTACCTGCCCAGCGTGTCTGCTAGA   720

Qy   721  CTCTATTTCCTCTGGACGGTGCCCCTGAGTGCCTGCGATCCCTCGTCCATGGGGCCCTG   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CTCTATTTCCTCTGGACGGTGCCCCTGAGTGCCTGCGATCCCTCGTCCATGGGGCCCTG   780

Qy   781  CACATAACCCTCAGCTTCACCTACAGCAGCATGCTGCGATCCCTGGGTGTATTATTT   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CACATAACCCTCAGCTTCACCTACAGCAGCATGCTGCGATCCCTGGGTGTATTATTT   840

Qy   841  TCAAGCCCCTCCTTTCCCAAATTCACACAACAAGCTCAAAATCTGAGTCTCAAAACCAAG   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TCAAGCCCCTCCTTTCCCAAATTCACACAACAAGCTCAAAATCTGAGTCTCAAAACCAAG   900

Qy   901  CAGCCAGGAGCACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTGCGC  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CAGCCAGGAGCACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTGCGC  960

Qy   961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC  1020

Qy  1021  CACATTGTTGAGTGGCAC  1038
          ||||||||||||||||||
Db  1021  CACATTGTTGAGTGGCAC  1038
```

RESULT 10
US-10-321-807-23
; Sequence 23, Application US/10321807
; Publication No. US20030166148A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Rupong
; APPLICANT: Darg, Huong T.
; APPLICANT: Lowitz, Kevin P.
; TITLE OF INVENTION: No. US20030166148A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN0086
; CURRENT APPLICATION NUMBER: US/10/321,807
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US/09/714,008
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,088
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,099
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,369
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/171,902
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,901
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,900
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/181,749
; PRIOR FILING DATE: 2000-02-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-321-807-23

Query Match   100.0%; Score 1038; DB 15; Length 1041;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1  ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACCACCCATCTCCCAGGTGATGCCGCCG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACCACCCATCTCCCAGGTGATGCCGCCG   60

Qy    61  CTGCTCATTGTGGCCTTTGTGTCTGGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CTGCTCATTGTGGCCTTTGTGTCTGGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC  120

Qy   121  TGCTTCCACATGAAGACCTGAAGCCAGCACTGTTTACCTTTTGGCCGTGGCT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  TGCTTCCACATGAAGACCTGAAGCCAGCACTGTTTACCTTTTGGCCGTGGCT  180

Qy   181  GATTTCCTCCTTATGATCTGCCTTTTCGGACGAGACTATTACCTCAGACGTAGACAC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GATTTCCTCCTTATGATCTGCCTTTTCGGACGAGACTATTACCTCAGACGTAGACAC  240
```

RESULT 11
US-10-210-172-177
; Sequence 177, Application US/10210172
; Publication No. US20040043928A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol
; APPLICANT: Rieger, Daniel
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Li, Li
; APPLICANT: Ji, Weizhen
; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Casman, Stacie
; APPLICANT: Voss, Edward
; APPLICANT: Boldog, Ferenc
; APPLICANT: Gorman, Linda
; APPLICANT: Leite, Mario
; APPLICANT: Vernet, Corine
; APPLICANT: Anderson, David
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zhong, Mei
; APPLICANT: Gerlach, Valerie
; APPLICANT: Hjalt, Tord
; APPLICANT: Rastelli, Luca
; APPLICANT: Spytek, Kimberly
; APPLICANT: Edinger, Shlomit
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Alsobrook II, John

SEQ ID NO 177
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)..(1043)
US-10-210-172-177

Query Match   100.0%; Score 1038; DB 13; Length 1050;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1038; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ATGTACAACGGGTCGTGCTGCGCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6 ATGTACAACGGGTCGTGCTGCGCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 65

Qy     61 CTGCTCATTGTGGCCCTTGTGCTGGGCGCACTAGGCGTGATGGGGTCCGCCTGTGTGGTTC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 CTGCTCATTGTGGCCCTTGTGCTGGGCGCACTAGGCGTGATGGGGTCCGCCTGTGTGGTTC 125

Qy    121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 185

Qy    181 GATTTCCTCCTTATGATCTGGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 GATTTCCTCCTTATGATCTGGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 245

Qy    241 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGGTGGCCATGAACAGGGCC 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGGTGGCCATGAACAGGGCC 305

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGACAGGTATTCAAAGTGGTCCACCCG 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGACAGTATTCAAAGTGGTCCACCCC 365

Qy    361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTGTCGCACCCTGTG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTGTCGCACCCTGTG 425

Qy    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACAATCCTCGCGTGCAAGAG 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACAATCCTCGCGTGCAAGAG 485

Qy    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGCTGCTGCATGACATCATG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGCTGCTGCATGACATCATG 545

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCCATCATCTTATTTTGCTCCTTCAAGATTGTT 605

Qy    601 TGGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGAGCTCGGATGAAGAGGCGACCCGG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 TGGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGCTCGGATGAAGAGGCGACCCCG 665

Qy    661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGGCTCACCTGCCCAGCGTGTCTGCTAGA 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGGCTCACCTGCCCAGCGTGTCTGCTAGA 725

Qy    721 CTCTATTTCCTCGGACGGTGCCCTGAGTGCCTGCGATCCCTCGTCTGTCCATGGGGCCCTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 CTCTATTTCCTCGGACGGTGCCCTGAGTGCCTGCGATCCCTCGTCTGTCCATGGGGCCCTG 785

Qy    781 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCTGGTGTATTATTTT 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCTGGTGTATTATTTT 845

Qy    841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAGCTCAAAATCTGAAAACCCAAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    846 TCAAGCCCCTCCTTTCCCAAATTCTACAACAGCTCAAAATCTGAAAACCCAAG 905

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGGTCGC 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGGTCGC 965

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGCATGGCATGGGATCCC 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    966 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGCATGGCATGGGATCCC 1025

Qy   1021 CACATTGTTGAGTGGCAC 1038
            ||||||||||||||||||
Db   1026 CACATTGTTGAGTGGCAC 1043
```

RESULT 12
US-10-044-643-1
; Sequence 1, Application US/10044643
; Publication No. US2003019535A1
; GENERAL INFORMATION:
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie J
; APPLICANT: Wolenc, Adam R
; APPLICANT: Spaderna, Steven K
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishnu, Vishun S
; APPLICANT: Tcherney, Velizar T

```
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Li, Li
;  APPLICANT:  Baumgartner, Jason C
;  APPLICANT:  Gusev, Vladimir
;  TITLE OF INVENTION: No. US2003019535A1el Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-748
;  CURRENT APPLICATION NUMBER: US/10/044,643
;  CURRENT FILING DATE: 2002-01-11
;  PRIOR APPLICATION NUMBER: 60/193,664
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 60/194,614
;  PRIOR FILING DATE: 2000-04-05
;  PRIOR APPLICATION NUMBER: 60/195,063
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: 60/195,066
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: 60/195,067
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: 60/195,068
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: 60/195,069
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: 60/195,070
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: 60/195,510
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: 60/219,855
;  PRIOR FILING DATE: 2000-07-21
;  PRIOR APPLICATION NUMBER: 60/221,284
;  PRIOR FILING DATE: 2000-07-27
;  PRIOR APPLICATION NUMBER: 60/221,325
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/224,588
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: 60/239,613
;  PRICR FILING DATE: 2000-10-11
;  PRIOR APPLICATION NUMBER: 60/262,508
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/263,604
;  PRIOR FILING DATE: 2001-01-23
;  PRIOR APPLICATION NUMBER: 60/263,433
;  PRIOR FILING DATE: 2001-01-23
;  PRIOR APPLICATION NUMBER: 60/265,161
;  PRIOR FILING DATE: 2001-01-30
;  NUMBER OF SEQ ID NOS: 83
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 1
;  LENGTH: 1050
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-044-643-1

Query Match           100.0%;  Score 1038;  DB 15;  Length 1050;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1038;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1  ATGTACAAACGGGTGCGCTGCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6  ATGTACAAACGGGTGCGCTGCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG   65

Qy    61  CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCGCACACCATCTCCCAGGTGTGGTTTC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66  CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCGCACACCATCTCCCAGGTGTGGTTTC  125

Qy   121  TGCTTCCACATGAAGACCTGGAAGCGAAGCCACACTGTTTACCTTTTCAATTGGCCGTGGCT  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126  TGCTTCCACATGAAGACCTGGAAGCCTGGAAGCCAGCACTGTTTACCTTTTCAATTGGCCGTGGCT  185

Qy   181  GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186  GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  245

Qy   241  TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACAGGGCC  300
```

```
Db   246  TGGGCTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACAGGGCC  305

Qy   301  GGGAGCACGTGTTCCTTACGGTGGTGGCTGCGGCAGGTATTTCAAAGTGTCCACCCC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGCAGGTATTTCAAAGTGTCCACCCC  365

Qy   361  CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGGCATCGTCTGCACCCTGTGG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366  CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGGCATCGTCTGCACCCTGTGG  425

Qy   421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG  485

Qy   481  ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTGTCGGCCAATGGCTGGCATGACATCATG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486  ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTGTCGGCCAATGGCTGGCATGACATCATG  545

Qy   541  TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546  TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  605

Qy   601  TGGAGCCTGAGGCGGAGGCGGAGGCAGCTGGCCAAGACAGCTGGCAGGCTCGGATGAAGAAGGCGACCCGG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606  TGGAGCCTGAGGCGGAGGCGGAGGCAGCTGGCCAAGACAGCTGGCAGGCTCGGATGAAGAAGGCGACCCGG  665

Qy   661  TTCATCATGTGGTGCAATTGTGTTCATCACAATGCTCACCTGCCCAGGTGTGTCTGCTAGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666  TTCATCATGTGGTGCAATTGTGTTCATCACAATGCTCACCTGCCCAGGTGTGTCTGCTAGA  725

Qy   721  CTCTATTTCCTCTGGACACGGTGCCCTGGAGTGCCTGCGGATCCCTGTGTCCATGGGGCCCTG  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726  CTCTATTTCCTCTGGACACGGTGCCCTGGAGTGCCTGCGGATCCCTGTGTCCATGGGGCCCTG  785

Qy   781  CACATAACCCTCAGCTTCACTACACAGTGAACAGCATGCTGGATCCCCTGGTGTATTATTT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786  CACATAACCCTCAGCTTCACTACACAGTGAACAGCATGCTGGATCCCCTGGTGTATTATTT  845

Qy   841  TCAAGCCCCTCCTTTCCCAAATTCTACACAACAAGCTCAAAATCTGAAACCCAAG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846  TCAAGCCCCTCCTTTCCCAAATTCTACACAACAAGCTCAAAATCTGAAACCCAAG  905

Qy   901  CAGGCCAGGACACTCAAAAAACACAAAGGCGGAAGAGATGCCAATTTCGAACCTCGGTCGC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906  CAGCCAGGACACTCAAAAAACACAAAGGCGGAAGAGATGCCAATTTCGAACCTCGGTCGC  965

Qy   961  AGGAGTTGCATCAGTCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966  AGGAGTTGCATCAGTCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1025

Qy  1021  CACATTGTTGAGTGGCAC  1038
          ||||||||||||||||||
Db  1026  CACATTGTTGAGTGGCAC  1043
```

RESULT 13
US-10-044-643-3
; Sequence 3, Application US/10044643
; Publication No. US2003019535A1
; GENERAL INFORMATION:
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie J
; APPLICANT: Wolenc, Adam R
; APPLICANT: Spaderna, Steven K
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishnu, Vishun S
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Gusev, Vladimir

TITLE OF INVENTION: No. US20030195335Ale1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-748
CURRENT APPLICATION NUMBER: US/10/044,643
CURRENT FILING DATE: 2002-01-11
PRIOR APPLICATION NUMBER: 60/193,664
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 60/194,614
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: 60/195,063
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,066
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,067
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,068
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,069
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,070
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,510
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/219,855
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: 60/221,284
PRIOR FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: 60/221,325
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/224,588
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/239,613
PRIOR FILING DATE: 2000-10-11
PRIOR APPLICATION NUMBER: 60/262,508
PRIOR FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: 60/263,433
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,604
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/265,161
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
  LENGTH: 1050
  TYPE: DNA
  ORGANISM: Homo sapiens
US-10-044-643-3

Query Match   100.0%; Score 1038; DB 15; Length 1050;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 65

Qy   61 CGCTCATTGTGCCTTGTGCTGGGCACTAGGCACACTGTTTACCTTTGTGGTTTC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 CTGCTCATTGTGCCTTGTGCTGGGCACTAGGCACACTGTTTACCTTTGTGGTTTC 125

Qy 121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 185

Qy 181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACACAGACTATTACCTTCAGACGTAGACAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 186 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACACAGACTATTACCTTCAGACGTAGACAC 245

Qy 241 TGGGCTTTTGGGGACATTCCCTCGCCGAGTGGGCGTCCTTCACGTTGGCCATGAACAGGGCC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 246 TGGGCTTTTGGGGACATTCCCTCGCCGAGTGGGCGTCCTTCACGTTGGCCATGAACAGGGCC 305

Qy 301 GGGAGCATCGTGTTCCTTACGGTGGTGCTGCGGCAGAGGTATTCAAAGTGGTCCACCCC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 306 GGGAGCATCGTGTTCCTTACGGTGGTGCTGCGGCAGAGGTATTCAAAGTGGTCCACCCC 365

Qy 361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 366 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTG 425

Qy 421 GCCCTGGTCATCCTGGAACACAGTGTATCTTTGCTGAGAACCATCCTGCGTGCAAGAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 426 GCCCTGGTCATCCTGGAACACAGTGTATCTTTGCTGAGAACCATCCTGCGTGCAAGAG 485

Qy 481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGCCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486 ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGCCATGACATCATG 545

Qy 541 TTCCAGCTGGAGTTCTTTATGCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 546 TTCCAGCTGGAGTTCTTTATGCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 605

Qy 601 TGGAGCCTGAGGCGGAGGCAGCAGTCTGGCCAGACAGCTCCGATGAAGAAGCGACCCGG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 606 TGGAGCCTGAGGCGGAGGCAGCAGTCTGGCCAGACAGCTCCGATGAAGAAGCGACCCGG 665

Qy 661 TTCATCATGGTGGTGCCAATTGTGTTCATCACATGTGCTCCTGCCCAGCGTGTCTGCTAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 666 TTCATCATGGTGGTGCCAATTGTGTTCATCACATGTGCTCCTGCCCAGCGTGTCTGCTAGA 725

Qy 721 CTCTATTTCCTCTGGACGGTGCCCTGCGATCCCTCGTGTCCATGGGGCCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 726 CTCTATTTCCTCTGGACGGTGCCCTGCGATCCCTCGTGTCCATGGGGCCCTG 785

Qy 781 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCCTGGTGTATTATTTT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 786 CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCCTGGTGTATTATTTT 845

Qy 841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGAGTCTGAAACCCAAG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 846 TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGAGTCTGAAACCCAAG 905

Qy 901 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 906 CAGCCAGGACACTCAAAAACACAAAGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 965

Qy 961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 966 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTGATGGGCAATGGGATCCC 1025

Qy 1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db 1026 CACATTGTTGAGTGGCAC 1043

RESULT 14
US-10-210-172-183
; Sequence 183, Application US/10210172
; Publication No. US2004004392BA1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol
; APPLICANT: Rieger, Daniel
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Li, Li
; APPLICANT: Ji, Weizhen
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Casman, Stacie
; APPLICANT: Voss, Edward
; APPLICANT: Boldog, Ferenc
; APPLICANT: Gorman, Linda
; APPLICANT: Leite, Mario
; APPLICANT: Vernet, Corine

```
;  APPLICANT:  Anderson, David
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Zhorg, Mei
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Edinger, Shlomit
;  APPLICANT:  Elleman, Karen
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-416 A
;  CURRENT APPLICATION NUMBER: US/10/210,172
;  CURRENT FILING DATE:  2001-08-01
;  PRIOR APPLICATION NUMBER: 60/309,501
;  PRIOR FILING DATE:  2001-08-02
;  PRIOR APPLICATION NUMBER: 60/323,994
;  PRIOR FILING DATE:  2001-09-21
;  PRIOR APPLICATION NUMBER: 60/373,814
;  PRIOR FILING DATE:  2002-04-19
;  PRIOR APPLICATION NUMBER: 60/310,291
;  PRIOR FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: 60/310,951
;  PRIOR FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: 60/310,544
;  PRIOR FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: 60/311,292
;  PRIOR FILING DATE:  2001-08-09
;  PRIOR APPLICATION NUMBER: 60/311,979
;  PRIOR FILING DATE:  2001-08-13
;  PRIOR APPLICATION NUMBER: 60/311,201
;  PRIOR FILING DATE:  2001-08-17
;  PRIOR APPLICATION NUMBER: 60/312,892
;  PRIOR FILING DATE:  2001-08-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 327
;  SOFTWARE: CuraSeqList version 0.1
;
; SEQ ID NO 183
;  LENGTH: 1057
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY:  CDS
;  LOCATION: (11)..(1048)
US-10-210-172-183
```

```
Query Match         100.0%;   Score 1038;   DB 13;   Length 1057;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 1038;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy       1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGGTGATGCCGCCG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGACACCATCTCCCAGGTGATGCCGCCG   70

Qy      61 CTGCTCCATTGTGGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGGTTTC  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 CTGCTCCATTGTGGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGGTTTC  130

Qy     121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT  190

Qy     181 GATTTCCTCCTTATGATCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC      240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  250

Qy     241 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGCCTCTTCACGTTGGCCATGAACAGGGCC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 TGGGCTTTGGGGACATTCCCTGCCGAGTGGGCCTCTTCACGTTGGCCATGAACAGGGCC  310
```

```
Qy     301 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 GGGAGCATCGTGTTCCTTACGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC  370

Qy     361 CACCACGCGGTGAACACTATCTCCACCGGTGGCGCTGGCATCGTCTGCACCCTGTGG  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 CACCACGCGGTGAACACTATCTCCACCGGTGGCGCTGGCATCGTCTGCACCCTGTGG  430

Qy     421 GCCCTGGTCATCCTGGGACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 GCCCTGGTCATCCTGGGACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  490

Qy     481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG  550

Qy     541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTGCTCCTTCAAGATTGTT  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTGCTCCTTCAAGATTGTT  610

Qy     601 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGTCGGATGAAGAAGGCGACCCGG  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     611 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGTCGGATGAAGAAGGCGACCCGG  670

Qy     661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     671 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA  730

Qy     721 CTCTATTTCCTCTGGACGGTGCGCCTCGAGTGCGTGCCTGCGGATCCCTCGTGTCATGGGGCCCTG  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     731 CTCTATTTCCTCTGGACGGTGCGCCTCGAGTGCGTGCCTGCGGATCCCTCGTGTCATGGGGCCCTG  790

Qy     781 CACATAACCCTCAGCTTCACTACATGAACAGCATGCTGGATCCCCTGGTGTATTATTT  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     791 CACATAACCCTCAGCTTCACTACATGAACAGCATGCTGGATCCCCTGGTGTATTATTT  850

Qy     841 TCAAGCCCCTCCTTTCCCAAATTCTACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     851 TCAAGCCCCTCCTTTCCCAAATTCTACAACAGCTCAAAATCTGCAGTCTGAAACCCAAG  910

Qy     901 CAGCCAGGACACTCAAAAACACAAAGGCGGAGAGATGCCAATTTCGAACCTCGGTCGC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     911 CAGCCAGGACACTCAAAAACACAAAGGCGGAGAGATGCCAATTTCGAACCTCGGTCGC  970

Qy     961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     971 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGATCCC  1030

Qy    1021 CACATTGTTGAGTGGCAC  1038
           ||||||||||||||||||
Db    1031 CACATTGTTGAGTGGCAC  1048
```

```
RESULT 15
US-10-210-172-191
; Sequence 191, Application US/10210172
; Publication No. US20040043928A1
; GENERAL INFORMATION:
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Li, Li
;  APPLICANT:  Ji, Weizhen
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Voss, Edward
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Gorman, Linda
```

```
;  APPLICANT:  Leite, Mario
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Anderson, David
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Zhong, Mei
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Edinger, Shlomit
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21432-416 A
;  CURRENT APPLICATION NUMBER: US/10/210,172
;  CURRENT FILING DATE:  2001-08-01
;  PRIOR APPLICATION NUMBER: 60/309,501
;  PRIOR FILING DATE:  2001-08-02
;  PRIOR APPLICATION NUMBER: 60/323,994
;  PRIOR FILING DATE:  2001-09-21
;  PRIOR APPLICATION NUMBER: 60/373,814
;  PRIOR FILING DATE:  2002-04-19
;  PRIOR APPLICATION NUMBER: 60/310,291
;  PRIOR FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: 60/310,951
;  PRIOR FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: 60/310,544
;  PRIOR FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: 60/311,292
;  PRIOR FILING DATE:  2001-08-09
;  PRIOR APPLICATION NUMBER: 60/311,979
;  PRIOR FILING DATE:  2001-08-13
;  PRIOR APPLICATION NUMBER: 60/313,201
;  PRIOR FILING DATE:  2001-08-17
;  PRIOR APPLICATION NUMBER: 60/312,892
;  PRIOR FILING DATE:  2001-08-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 327
;  SOFTWARE: CuraSeqlist version 0.1
;
;  SEQ ID NO 191
;  LENGTH: 1057
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (2)..(1057)
US-10-210-172-191
```

Query Match   100.0%;   Score 1038;   DB 13;   Length 1057;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     1 ATGTACAACGGGTGCGTCGCCGCATCGAGAGGGGACACCATCTCCCAGGTGATGCCGGCCG    60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    11 ATGTACAACGGGTCGGTGCGCTGCGCGCATCGAGAGGGGACACCATCTCCCAGGTGATGCCGGCCG    70

Qy    61 CTGCTCATTGTGGCCTTGTGCTCGGCGGCACTAGGCAATGGGGTCGCCCTGTGGGTTTC    120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71 CTGCTCATTGTGGCCTTGTGCTCGGCGGCACTAGGCAATGGGGTCGCCCTGTGTGGGTTTC    130

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACACTGTTTACCTTTCAATTGGCCGTGGCT    180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 TGCTTCCACATGAAGACCTGGAAGCCCAGCACACTGTTTACCTTTCAATTGGCCGTGGCT    190

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACGACTATTACCTCAGACGTAGACAC    240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC    250

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC    300
```

```
Db   251 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGGCTCTTCACGTTGGCCATGAACAGGGCC    310

Qy   301 GGGAGCATCGTTCCTTACGGTGGTGGCTGCGGCAGGTATTTCAAAGTGTCCACCCC    360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGCAGGTATTTCAAAGTGTCCACCCC    370

Qy   361 CACCACGCGGTGAACTATTCCACCCGGTGGCGGCTGGCCATCGTCCTGCACCCTGTGG    420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 CACCACGCGGTGAACTATCTCCACCCGGTGGCGGCTGGCCATCGTCCTGCACCCTGTGG    430

Qy   421 GCCCTGGTCATCCTCGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGGCGTGCAAGAG    480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 GCCCTGGTCATCCTCGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGGCGTGCAAGAG    490

Qy   481 ACGGCCGTCTCCTGTACATGGAGCTTCATCATGGAGTCGGCCAATGGCTGGCCATGACATCATG    540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 ACGGCCGTCTCCTGTCCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCCATGACATCATG    550

Qy   541 TTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTATTTTGCTCCTTCAAGATTGTT    600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   551 TTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTATTTTGCTCCTTCAAGATTGTT    610

Qy   601 TGGAGCCTGGAGGCGAGCAGCTGGCCAGACGAGCTCGGATGAAGAAGGCGACCCGG    660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   611 TGGAGCCTGGAGGCGAGCAGCTGGCCAGACGAGCTCGGATGAAGAAGGCGACCCGG    670

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA    720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   671 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA    730

Qy   721 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGGCCTG    780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   731 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGGCCTG    790

Qy   781 CACATAACCCTGCAGCTTCACCTACATGAACAGAGCTGGATCCCCTGGTGTTGTATTATTT    840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   791 CACATAACCCTGCAGCTTCACCTACATGAACAGAGCTGGATCCCCTGGTGTTGTATTATTT    850

Qy   841 TCAAGCCCCTTCCTTTCCCAAATCTACAACAAGTTCAAAATCTGCAGTCTGAAACCCAAG    900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   851 TCAAGCCCCTTCCTTTCCCAAATCTACAACAAGTTCAAAATCTGCAGTCTGAAACCCAAG    910

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC    960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   911 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC    970

Qy   961 AGGAGTTGGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC    1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   971 AGGAGTTGGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC    1030

Qy  1021 CACATTGTTGAGTGGCAC    1038
           ||||||||||||||||||
Db  1031 CACATTGTTGAGTGGCAC    1048
```

Search completed: July 3, 2004, 10:07:49
Job time : 565 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        June 30, 2004, 17:15:22 ; Search time 17 Seconds
                                         (without alignments)
                                  1059.781 Million cell updates/sec

Title:        US-10-076-260-2
Perfect score: 1853
Sequence:      1 MYNGSCCRIEGDTISQVMPP.........ANSFQSQSDGQWDPHIVEWH 346

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:    141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES
```

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1853	100.0	346	1	GP81_HUMAN	Q9bxc0 homo sapien
2	1483	80.0	343	1	GP81_MOUSE	Q8c131 mus musculu
3	880.5	47.5	387	1	HM74_HUMAN	P49019 homo sapien
4	448.5	24.2	319	1	GP31_HUMAN	O00270 homo sapien
5	371.5	20.0	309	1	CLT2_MOUSE	Q920a1 mus musculu
6	366.5	19.8	346	1	CLT2_HUMAN	Q9ns75 homo sapien
7	364	19.6	362	1	P2YR_CHICK	P34996 gallus gall
8	364	19.6	362	1	P2YR_MELGA	P49652 meleagris g
9	362.5	19.6	373	1	P2YR_BOVIN	P48042 bos taurus
10	362	19.5	345	1	CLT2_PIG	Q95n03 sus scrofa
11	358.5	19.3	373	1	P2YR_RAT	P49651 rattus norv
12	356.5	19.2	373	1	P2YR_CAVPO	P59902 cavia porce
13	355.5	19.2	309	1	CLT2_RAT	Q924t9 rattus norv
14	354	19.1	373	1	P2YR_MOUSE	P47900 mus musculu
15	351.5	19.0	373	1	P2YR_HUMAN	P49650 mus musculu
16	346.5	18.7	370	1	P2Y9_HUMAN	Q99677 homo sapien
17	340	18.3	344	1	P2Y5_MOUSE	Q8bmc0 mus musculu
18	339	18.3	308	1	P2Y5_CHICK	P32250 gallus gall
19	336	18.1	367	1	GP17_HUMAN	Q13304 homo sapien
20	327	17.6	309	1	GP35_HUMAN	Q9hc97 homo sapien
21	325	17.5	344	1	P2Y5_HUMAN	P43657 homo sapien
22	320	17.3	373	1	P2Y2_MOUSE	P35383 mus musculu
23	316.5	17.1	374	1	P2Y2_RAT	P41232 rattus norv
24	310.5	16.8	420	1	PAR1_XENLA	P47749 xenopus lae
25	309.5	16.7	363	1	SSR5_RAT	P30938 rattus norv
26	308.5	16.6	307	1	GP35_MOUSE	Q9es90 mus musculu
27	308	16.6	340	1	CLT1_PIG	Q95n02 sus scrofa
28	307	16.6	391	1	SSR1_MOUSE	P30873 mus musculu
29	305.5	16.5	391	1	SSR1_HUMAN	P30872 homo sapien
30	305.5	16.5	391	1	SSR1_RAT	P28646 rattus norv
31	305	16.5	337	1	CLT1_HUMAN	Q9y271 homo sapien
32	305	16.5	372	1	GP92_HUMAN	Q9h1c0 homo sapien
33	305	16.5	377	1	P2Y2_HUMAN	P41231 homo sapien
34	304.5	16.4	359	1	AG2R_BOVIN	P25104 bos taurus
35	302.5	16.3	369	1	SSR2_HUMAN	P30874 homo sapien
36	302	16.3	391	1	BRB2_HUMAN	P30411 homo sapien
37	302	16.3	399	1	PAR2_MOUSE	P55086 mus musculu
38	300.5	16.2	369	1	SSR2_MOUSE	P30875 mus musculu
39	300.5	16.2	369	1	SSR2_RAT	P30680 rattus norv
40	300	16.2	328	1	P2Y6_RAT	Q63371 rattus norv
41	300	16.2	384	1	SSR4_RAT	P30937 rattus norv
42	300	16.2	428	1	SSR3_MOUSE	P30935 mus musculu
43	299.5	16.2	369	1	SSR2_PIG	P34994 sus scrofa
44	299	16.1	388	1	SSR4_HUMAN	P31391 homo sapien
45	299	16.1	397	1	PAR2_RAT	Q63645 rattus norv

```
                                                        ALIGNMENTS

RESULT 1
GP81_HUMAN
ID   GP81_HUMAN     STANDARD;      PRT;    346 AA.
AC   Q9BXC0;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable G protein-coupled receptor GPR81.
GN   GPR81 OR FKSG80.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=21458557; PubMed=11574155;
RA   Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,
RA   Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RT   "Discovery and mapping of ten novel G protein-coupled receptor
RT   genes.";
RL   Gene 275:83-91(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22040266; PubMed=12044878;
RA   Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT   "Identification of G protein-coupled receptor genes from the human
RT   genome sequence.";
RL   FEBS Lett. 520:97-101(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Wang Y.-G., Gong L.;
RT   "Molecular cloning of FKSG80, a novel gene encoding a putative
RT   chemokine receptor.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA   Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT   "Genome-wide discovery and analysis of human seven transmembrane helix
RT   receptor genes.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in pituitary tissue. Not detected in
CC       frontal, temporal and occipital lobes of the cortex, basal
CC       forebrain, caudate nucleus, nucleus accumbens, and hippocampus.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
```

```
DR   EMBL; AF411110; AAL26481.1; -.
DR   EMBL; AB083631; BAB89344.1; -.
DR   EMBL; AF345568; AAK29071.1; -.
DR   EMBL; AB065866; BAC06084.1; -.
DR   Genew; HGNC:4532; GPR81.
DR   MIM; 606923; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     21       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     22     42       1 (POTENTIAL).
FT   DOMAIN       43     49       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     50     70       2 (POTENTIAL).
FT   DOMAIN       71     89       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     90    110       3 (POTENTIAL).
FT   DOMAIN      111    130       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    131    151       4 (POTENTIAL).
FT   DOMAIN      152    182       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    183    203       5 (POTENTIAL).
FT   DOMAIN      204    220       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    221    241       6 (POTENTIAL).
FT   DOMAIN      242    261       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    262    281       7 (POTENTIAL).
FT   DOMAIN      282    346       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     88      3       BY SIMILARITY.
FT   CARBOHYD      3      3       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   346 AA;  39295 MW;  B0DB14EEB3A47A5 CRC64;
```

```
Query Match             100.0%;  Score 1853; DB 1; Length 346;
Best Local Similarity   100.0%;  Pred. No. 1e-125;
Matches  346;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCHMKTWKPSTVYLFNLAVA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCHMKTWKPSTVYLFNLAVA 60

Qy         61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHP 120

Qy        121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSRSFIMRSANGWHDIM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCSRSFIMESANGWHDIM 180

Qy        181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240

Qy        241 LYFLWTVPSSACDPSVHGALFITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LYFLWTVPSSACDPSVHGALFITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300

Qy        301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH 346
```

```
RESULT 2
GP81_MOUSE
ID   GP81_MOUSE     STANDARD;     PRT;   343 AA.
AC   Q8C131;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable G protein-coupled receptor GPR81.
GN   GPR81.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Skin;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shinada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AK029064; BAC26273.1; -.
DR   MGD; MGI:2441671; Gpr81.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     21       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     22     42       1 (POTENTIAL).
FT   DOMAIN       43     49       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     50     70       2 (POTENTIAL).
FT   DOMAIN       71     90       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     91    111       3 (POTENTIAL).
FT   DOMAIN      112    131       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    132    152       4 (POTENTIAL).
FT   DOMAIN      153    182       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    183    203       5 (POTENTIAL).
FT   DOMAIN      204    220       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    221    241       6 (POTENTIAL).
FT   DOMAIN      242    259       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    260    280       7 (POTENTIAL).
FT   DOMAIN      281    343       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     88    165       BY SIMILARITY.
FT   CARBOHYD      3      3       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   343 AA;  38927 MW;  917FA9499B2B03FD CRC64;

Query Match            80.0%;  Score 1483; DB 1; Length 343;
```

Best Local Similarity 81.4%; Pred. No. 2.8e-99;
Matches 276; Conservative 19; Mismatches 44; Indels 0; Gaps 0;

```
Qy    1 MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVA 60
          ||| ||||||||| |:.||||||||| :|| |||||||||||||||||||||| ||||||||
Db    1 MDNGSCCLIEGRPISQVMPPLLILVFVLGALGNGIALCGFCFHMKTWKSSTIYLFNLAVA 60

Qy   61 DFLLMICLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVAADRYFKVVHP 120
          |||||||||| ||||||||||| |  |||||||||| |||||||||| |||||||||||
Db   61 DFLLMICLPLRTDYLRRRHWIFGDIACRIVLFKLAMNRAGSIVFLTVAVDRYFKVVHP 120

Qy  121 HHAVNTISTRVAAGIVCLWALVILGTVYLLLLENHLCVQETAVSCESFIMESANGWHDIM 180
          ||:|| ||||||||||| :|||: |||| ||||:||||||||||||||||||||||||||
Db  121 HHMVNAISNRTAAATACVLWTLVILGTVYLLMESHLCVQGTLSSCESFIMESANGWHDVM 180

Qy  181 FQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
          |||| |||||||||||| ||:||||||||| :|| | |||||||||:||||||| |||||
Db  181 FQLEFFLPLTIILFCSVNVVWSLRRRQQLTRQARMRRAATRFIMVVASVFITCYLPSVLAR 240

Qy  241 LYFLWTVPSSACDPSVHGALHITLSFFTYMNSMLDPLWYFSSPSFPKFYNKLKICSLKPK 300
          |||||||| :||||||| ||:|||:|||||||||||||||| :||||||||| ||||||||
Db  241 LYFLWTVFTSACDPSVHTALHVTLSFTYLNSMLDPLVYFSSPSLPKFYTKLITCISLKPK 300

Qy  301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWD 339
          || :||||||||||||| |:|||   ||| |||| |||||
Db  301 RPGRTKTRRSEEMPISNLCSKSSIDGANRSQRPSDGQWD 339
```

KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN   1   28   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   29   50   1 (POTENTIAL).
FT   DOMAIN   51   63   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   64   85   2 (POTENTIAL).
FT   DOMAIN   86   102   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   103   123   3 (POTENTIAL).
FT   DOMAIN   124   142   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   143   163   4 (POTENTIAL).
FT   DOMAIN   164   194   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   195   209   5 (POTENTIAL).
FT   DOMAIN   210   236   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   237   256   6 (POTENTIAL).
FT   DOMAIN   257   273   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   274   298   7 (POTENTIAL).
FT   DOMAIN   299   387   CYTOPLASMIC (POTENTIAL).
FT   DISULFID   100   177   BY SIMILARITY.
SQ   SEQUENCE   387 AA; 44481 MW; C244F562C2343647 CRC64;

Query Match   47.5%;   Score 880.5; DB 1; Length 387;
Best Local Similarity 52.2%; Pred. No. 3.8e-56;
Matches 178; Conservative 49; Mismatches 107; Indels 7; Gaps 4;

```
Qy    5 SCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLL 64
          .||| |||||||| | :||::||:|:|||||| |:|::||||||||:::||||||||||
Db   17 NCCVFRDDFIAKVLPPVGLREFIFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVADFLL 76

Qy   65 MICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHAV 124
          :|| ||| |  ||||:||||||||||:|||:|:|||||::||||| ||||||||||:||||
Db   77 IICLPFVMDYVVRRSDWNFGDIPCRLVLFMFAMNRQGSIIFJTVVAVDRYFRVVHPHHAL 136

Qy  125 NTISTRVAAGIVCTLWALVILGTVYLLLLENHLCVQETAVSCESFIMESANGWHDIMFQLE 184
          |.|| |||||:| :|:|||:|||| :|:||| ||| :|  ||||||||:||||||||||:||||
Db  137 NKISNWTAAIISCLLWGITVGLTVHLLKKKLLIQNGPANVCISFSICHTFRWHEAMFLLE 196

Qy  185 FFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSARLYFL 244
          || :||||:||||||||||||||| :||| |||:||:||:|||||||||:||||| |||||
Db  197 FLLPLGIILFCSARIIWSLRQR-QMDRHAKIKRAITFIMVVAIVFVICFLPSVVRIRIF 255

Qy  245 WTVPSSA---CD--PSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKP 299
          ||::|:|    :|    | :|| :|:|| ||:|||||:||::|::|||||| |:|||:|||||
Db  256 WLLHTSGTQNCEVYRSVDLAFFITLSFTYMNSMLDPVVYYFSSSPSFPNFFSTLINRCLQR 315

Qy  300 KQPGHSKTQRPBEMPISNLGRRSCISVANSFQSQSDGQWDP 340
          | :|:::|::      ::    :||::  ||
Db  316 KMTGEPDNNRSTSVELTGDPNKT-RGAPEAIMANSGEPWSP 355
```

RESULT 3
HM74_HUMAN
ID   HM74_HUMAN   STANDARD;   PRT;   387 AA.
AC   P49019;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable G protein-coupled receptor HM74.
GN   HM74.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Monocytes;
RX   MEDLINE=94092629; PubMed=7505609;
RA   Nomura H., Nielsen B.W., Matsushima K.;
RT   "Molecular cloning of cDNAs encoding a LD78 receptor and putative
RT   leukocyte chemotactic peptide receptors.";
RL   Int. Immunol. 5:1239-1249(1993).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; D10923; BAA01721.1; -.
DR   PIR; I69202; I69202.
DR   HSSP; P34996; 1DED.
DR   MIM; 606039; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

RESULT 4
GP31_HUMAN
ID   GP31_HUMAN   STANDARD;   PRT;   319 AA.
AC   O00270;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable G protein-coupled receptor GPR31.
GN   GPR31.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=97349123; PubMed=9205127;
RA   Zingoni A., Rocchi M., Storlazzi C.T., Bernardini G., Santoni A.,
RA   Napolitano M.;
RT   "Isolation and chromosomal localization of GPR31, a human gene
RT   encoding a putative G protein-coupled receptor.";
RL   Genomics 42:519-523(1997).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

```
CC   -------------------------------------------------------------------------
```
```
CC   -------------------------------------------------------------------------
DR   EMBL; U65402; AAC51375.1; -.
DR   Genew; HGNC:4486; GPR31.
DR   MIM; 602043; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN        1     16    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     17     37    1 (POTENTIAL).
FT   DOMAIN       38     52    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     53     73    2 (POTENTIAL).
FT   DOMAIN       74     91    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     92    110    3 (POTENTIAL).
FT   DOMAIN      111    131    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    132    152    4 (POTENTIAL).
FT   DOMAIN      153    180    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    181    201    5 (POTENTIAL).
FT   DOMAIN      202    219    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    220    240    6 (POTENTIAL).
FT   DOMAIN      241    265    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    266    284    7 (POTENTIAL).
FT   DOMAIN      285    319    CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5    N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   319 AA;  35079 MW;  7A4F164CD9C969DD CRC64;

Query Match     24.2%;  Score 448.5;  DB 1;  Length 319;
Best Local Similarity   34.0%;  Pred. No. 2.4e-25;
Matches  100;  Conservative  62;  Mismatches 121;  Indels  11;  Gaps  5;

Qy    7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADPLLMI 66
Db    6 CSAPSTVVATAVGVLLGLECGLGLLGNAVALWTFLFRVRVWKPYAVVILNLALADLLLAA 65

Qy   67 CLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 126
Db   66 CLPFLAAFYLSLQAWHLGRVGCWALRFLLDLLSRSVGMAFLRAWALDRYLRVVHPRLKVNL 125

Qy  127 ISTRVAAGIVCTLWALLVILGTVYLLLENHLCVQETAVSCESPIMESANG----WHDIMF 181
Db  126 LSPQALGVSGLVWLLMVALTCPGLLISE-AAQNSTRCHSF-YSRADGSFSIIWQEALS 182

Qy  182 QLEFFMPLGIILFCSFKIVWSLRRR-QQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
Db  183 CLQFVLPFGLIVFCNAGIIRALQKRLREPEKQPKLQRAQALVTLVVVLFALCFLPCFLAR 242

Qy  241 --LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSPSFKFYNKL 292
Db  243 VLMHIFQNLGSCRALCAVAHTSDVTGSLTYLHSVVNPVVYCFSSPTFRSSVYRRV 296
```

RESULT 5

CLT2_MOUSE

```
ID   CLT2_MOUSE      STANDARD;      PRT;   309 AA.
AC   Q920A1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 2 (CysLTR2).
GN   CYSLTR2 OR CYSLT2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=21601669; PubMed=11591709;
RA   Hui Y., Yang G., Galczenski H., Figueroa D.J., Austin C.P.,
RA   Copeland N.G., Gilbert D.J., Jenkins N.A., Funk C.D.,
RT   "The murine cysteinyl leukotriene 2 (CysLT2) receptor. cDNA and
RT   genomic cloning, alternative splicing, and in vitro
RT   characterization.";
RL   J. Biol. Chem. 276:47489-47495(2001).
CC   -!- FUNCTION: Receptor for cysteinyl leukotriene. The response is
CC       mediated via a G-protein that activates a phosphatidylinositol-
CC       calcium second messenger system. The rank order of affinities for
CC       the leukotrienes is LTC4 = LTD4 >> LTE4.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Widely expressed at low levels, with highest
CC       expression in the spleen, thymus and adrenal gland, and lower in
CC       the kidney, brain and peripheral blood leukocytes.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -------------------------------------------------------------------------
```
```
CC   -------------------------------------------------------------------------
DR   EMBL; AF331658; AAK97354.1; -.
DR   MGD; MGI:1917336; Cysltr2.
DR   GO; GO:0016020; C:membrane; IDA.
DR   GO; GO:0001631; F:cysteinyl leukotriene receptor activity; IDA.
DR   InterPro; IPR004071; Cysleuk_receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF30001; 7tm_1; 1.
DR   PRINTS; PR01533; CYSLTRECPTR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     26    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     47    1 (POTENTIAL).
FT   DOMAIN       48     56    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     57     77    2 (POTENTIAL).
FT   DOMAIN       78     98    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     99    119    3 (POTENTIAL).
FT   DOMAIN      120    138    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    139    159    4 (POTENTIAL).
FT   DOMAIN      160    187    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    188    208    5 (POTENTIAL).
FT   DOMAIN      209    229    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    230    250    6 (POTENTIAL).
FT   DOMAIN      251    271    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    272    292    7 (POTENTIAL).
FT   DOMAIN      293    309    CYTOPLASMIC (POTENTIAL).
FT   DISULFID     95    171    BY SIMILARITY.
FT   CARBOHYD     14     14    N-LINKED (GLCNAC...). (POTENTIAL).
SQ   SEQUENCE   309 AA;  35227 MW;  327B14A6EDDD2A02 CRC64;

Query Match     20.0%;  Score 371.5;  DB 1;  Length 309;
Best Local Similarity   32.7%;  Pred. No. 7.5e-20;
Matches   96;  Conservative   56;  Mismatches 125;  Indels  17;  Gaps  7;

Qy    2 YNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPST---VYLFNLA 58
Db   10 YSNRNCTIE-NFKKEFYPIIYLIIFFWGALGNGFSIYVF---LQTCKKSTSVNVFMLNLA 65

Qy   59 VADFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVV 118
```

```
Db        66  TSDFLFISTLPFRADYYFRGSNWIFGDLACRVMSYSLYVNMYTSIYFLTVLSVRFQATV 125
               :| .| .:| . ||.|.|   | . :| .: | | : || :: |.: | | |.:|
Qy       119  HPHHAVNT-STRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSC-----ESFIMESA 173
               |.| ||:  .:| |:|.| || | :   :||| :  | .| :  |      | :||  .
Db       126  HPFRMFHVTSVR-SAWILCGIIWVFMASSALLLVNGQBEKDNIISCLELSPQKF--KSL 182
               | ..| .|   |   |:| ||:.:| |  |   .| : : | . . : .  :  |
Qy       174  KGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQ--QLARQARMKKATRPIMVVAIVFIT 231
               | :    | .| | . |: : |. ||. | : :       | :  : .| ||:|| | |
Db       183  LIWNHIAVAVGFLLPFLITUTVCYLLIRILLKARIPESGPRAAHRKALTTIVIAMITFLL 242
               | :.| . | |  .|: ||   | | |  .   |     || : . |.: |  |. :
Qy       232  CYLPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSF 285
               | .|:.| .||       |.   | | :  | : || : .  | |::|| ..|
Db       243  CFLPYHALRTLHLVTWDKDSCGDVLHKATVITLTMAAANSCFNPFLYYFAGENF 296
```

RESULT 6

```
ID   CLT2_HUMAN     STANDARD;      PRT;   346 AA.
AC   Q9NS75; Q9HCQ2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Cysteinyl leukotriene receptor 2 (CysLTR2) (PSEC0146) (HG57) (HPN321)
DE   (hGPCR21).
GN   CYSLTR2 OR CYSLT2 OR CYSLTR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=20374466; PubMed=10913337;
RA   Takasaki J., Kamohara M., Matsumoto M., Saito T., Sugimoto T., Ota T.,
RA   Nishikawa T., Kawai Y., Masuho Y., Isogai T., Suzuki Y., Sugano S.,
RA   Furuichi K.;
RT   "The molecular characterization and tissue distribution of the human
RT   cysteinyl leukotriene CysLT2 receptor.";
RL   Biochem. Biophys. Res. Commun. 274:316-322(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20459128; PubMed=10851239;
RA   Heise C.E., O'Dowd B.F., Figueroa D.J., Sawyer N., Nguyen T.,
RA   Im D.-S., Stocco R., Bellefeuille J.N., Abramovitz M., Cheng R.,
RA   Williams D.L. Jr., Zeng Z., Liu Q., Ma L., Clements M.K., Coulombe N.,
RA   Liu Y., Austin C.P., George S.R., O'Neill G.P., Metters K.M.,
RA   Lynch K.R., Evans J.F.;
RT   "Characterization of the human cysteinyl leukotriene 2 receptor.";
RL   J. Biol. Chem. 275:30531-30536(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20545741; PubMed=11093801;
RA   Nothacker H.-P., Wang Z., Zhu Y., Reinscheid R.K., Lin S.H.S.,
RA   Civelli O.;
RT   "Molecular cloning and characterization of a second human cysteinyl
RT   leukotriene receptor: discovery of a subtype selective agonist.";
RL   Mol. Pharmacol. 58:1601-1608(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Dunn M.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22040266; PubMed=12044878;
RA   Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT   "Identification of G protein-coupled receptor genes from the human
RT   genome sequence.";
RL   FEBS Lett. 520:97-101(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Kopatz S.A., Aronstam R.S., Sharma S.V.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
```

RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 17-346 FROM N.A.
RA   Suga H.;
RT   "Homo sapiens cysteinyl leukotriene receptor 1 like receptor.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.

```
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC       mediated via a G-protein that activates a phosphatidylinositol-
CC       calcium second messenger system. Stimulation by BAY u9773, a
CC       partial agonist, induces specific contractions of pulmonary veins
CC       and might also have an indirect role in the relaxation of the
CC       pulmonary vascular endothelium. The rank order of affinities for
CC       the leukotrienes is LTC4 = LTD4 >> LTE4.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in the
CC       heart, placenta, spleen, peripheral blood leukocytes and adrenal
CC       gland. In lung, expressed in the interstitial macrophages, and
CC       slightly in smooth muscle cells.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB038269; BAB03601.1; -.
DR   EMBL; AF254664; AAG17281.1; -.
DR   EMBL; AF279611; AAK69485.1; -.
DR   EMBL; AL137118; CAC29102.1; -.
DR   EMBL; AB083603; BABB9316.1; -.
DR   EMBL; AY389504; AAQ91330.1; -.
DR   EMBL; AB041644; BAB16379.1; -.
DR   Genew; HGNC:18274; CYSLTR2.
DR   MIM; 605666; -.
DR   GO; GO:0004974; F:leukotriene receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR004071; Cysleuk_receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01533; CYSLTRECPTR.
DR   PRINTS; PR00237; GPCRRHODOPSN
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN         1     42       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      43     63       1 (POTENTIAL).
FT   DOMAIN        64     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      73     93       2 (POTENTIAL).
FT   DOMAIN        94    123       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     124    144       3 (POTENTIAL).
FT   DOMAIN       145    153       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     154    174       4 (POTENTIAL).
FT   DOMAIN       175    204       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     205    225       5 (POTENTIAL).
FT   DOMAIN       226    245       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     246    266       6 (POTENTIAL).
FT   DOMAIN       267    286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     287    307       7 (POTENTIAL).
FT   DOMAIN       308    346       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     111    187       BY SIMILARITY.
FT   CARBOHYD      20     20       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      26     26       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      30     30       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     181    181       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   346 AA;  39635 MW;  EB54A4A2DDC5EE4 CRC64;
```

Query Match   19.8%; Score 366.5; DB 1; Length 346;
Best Local Similarity 29.9%; Pred. No. 1.9e-19;
Matches 94; Conservative 69; Mismatches 134; Indels 17; Gaps   8;

```
Qy      3 NGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPST---VYLFNLAV 59
            ::|:||  |  |||  :|   |||||||||||||||||||||    ::: ||::| |||
Db     27 MSRNCTIE-NFKREFFIVYLIIFFWGVLGNGLSIYVF---LQPYKKSTSVNVFMLNLAI 82

Qy     60 ADFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVH 119
            :||:||| |||:|||  ||||  |: ||:|| |: ||   :|:||: |: : .||:|
Db     83 SDLLFISTLPFFRADYYLRGSNWIFGDLACRIMSYSLYVNMYSSIYFLTVLSVVRFLAMVH 142

Qy    120 PHHAVNTISTRVAAGIVCTLWALIVLGTVYLLENHLCVQETAVSC---ESFIMESANGW 176
            | :| |: :||  :| | | ||: ||    |||||||||||    ||:|||||||
Db    143 PFRLLHVTSIR-SAWILCGIIWILIMASSIMLLDSGSEQNGSVTSCLELNLYKIAKLQTM 201

Qy    177 HDIMFQLEFFMPLGIILFCSFKIVWSLRRRQ--QLARQARMKKATRFIMVVAIVFITCYL 234
            :| ||   ||| :  :: ::|:||| | |     :     ||::| ||: ||:||||||
Db    202 NYIALVVGCLLPFFTLSICYLLIIRVLLKVEVPESGLRVSHRXALTTIIITLIIFFLCFL 261

Qy    235 PSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYPSSPSPFKFYNKLKI 294
            || ||||||| |:|||  |: ||:|:::|: ||::| |||||||| ||: ||  :|:
Db    262 PYHTLRTWHLTTWKVGLCKDRLHKALVTLALAAAANACFNPLLYYFAGENFK---DRLK- 317

Qy    295 CSLKPKQPGHSKTQ 308
            |||:|:||||:|
Db    318 SALRKGHPQKAKTK 331
```

RESULT 7
```
P2YR_CHICK
ID   P2YR_CHICK         STANDARD;       PRT;   362 AA.
AC   P34996;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=93285340; PubMed=8508924;
RA   Webb T.E., Simon J., Krishek B.J., Bateson A.N., Smart T.G.,
RA   King B.F., Burnstock G., Barnard E.A.;
RT   "Cloning and functional expression of a brain G-protein-coupled ATP
RT   receptor.";
RL   FEBS Lett. 324:219-225 (1993).
RN   [2]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=97026278; PubMed=8872457;
RA   van Rhee A.M., Fischer B., van Galen P.J.M., Jacobson K.A.;
RT   "Modelling the P2Y purinoceptor using rhodopsin as template.";
RL   Drug Des. Discov. 13:133-140(1995).
CC   -!- FUNCTION: Receptor for extracellular adenine nucleotides such as
CC       ATP and ADP. Seems to mediate its action via a pertussis toxin
CC       insensitive G-protein, probably belonging to the Gq family that
CC       activate a phosphatidylinositol-calcium second messenger system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Brain, spinal cord, gastrointestinal tract,
CC       spleen and leg muscle. Is not detected in the heart, liver,
CC       stomach, lung and kidney.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement.(See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X73268; CAA51716.1; -.
```

```
DR   PIR; S33733; S33733.
DR   PDB; 1DDD; 11-JUL-96.
DR   InterPro; IPR000276; GPCR_Rhodopsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; 3D-structure.
FT   DOMAIN        1     41       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     42     63       1 (POTENTIAL).
FT   DOMAIN       64     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     77     98       2 (POTENTIAL).
FT   DOMAIN       99    115       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    116    136       3 (POTENTIAL).
FT   DOMAIN      137    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    177       4 (POTENTIAL).
FT   DOMAIN      178    207       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    208    227       5 (POTENTIAL).
FT   DOMAIN      228    254       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    255    274       6 (POTENTIAL).
FT   DOMAIN      275    292       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    293    317       7 (POTENTIAL).
FT   DOMAIN      318    362       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    113    191       BY SIMILARITY.
FT   CARBOHYD     11     11       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    102    102       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC...) (POTENTIAL).
FT   HELIX        42     69
FT   HELIX        77    102
FT   HELIX       115    137
FT   HELIX       154    178
FT   HELIX       205    231
FT   HELIX       250    275
FT   TURN        290    305
FT   HELIX       306    307
FT   HELIX       308    320
SQ   SEQUENCE   362 AA;  41194 MW;  AB06C8BFB9514761 CRC64;
```

```
Query Match             19.6%;  Score 364;  DB 1;  Length 362;
Best Local Similarity   29.1%;         Pred. No. 3e-19;
Matches   95;  Conservative  58;  Mismatches 130;  Indels  44;  Gaps   7;

Qy     18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLR 77
           :|:|:|||:||| |||  |  ::: :::|||||||  :|||| ||::||::||::||:|
Db     43 LPTVYILVFITGFPLGNSVAIWMFVFHMRPWSGISVYMFNLALADFLYVLTLPALIFYYFN 102

Qy     78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
           | |:|||||:||:|||| | :|||  ||||||  ||::|: |:: :|:|| :| |:|
Db    103 KIDWIFGDVMCKLQRFIFHVNLYGSILFLTCISVHRYTGVVHPLIKSLGRLKKKNAVVYSS 162

Qy    138 TIWALVIL--------GTVYLLENHLCVQETA---------VSCESFIMESANGWHD 178
           |:||||:|         ||:||||||||||:           |:| |||||||||
Db    163 LVWALVVAVIAPILFYSGTGVRRNKITCYDTTADEYLRSYFVYSMCTTVFM------- 214

Qy    179 IMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP--- 235
           ::||::||:|||:|| || |  | ::::|:::|||:|  :|:|  |:|: |||| |
Db    215 -----RCIPFIVILGCYGLIVKALIYK-DLDNSPLRRKSIYLVIVILTVFAVSYLPFHV 267

Qy    236 ---SVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYPSSPSPFKFYNK 291
              | | ||||||| || :||||:| :|:||:||||||:||||||::  : : |  :||
Db    268 MKTLNLRARLDF-QTPQMCAFNDKVXYATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR 326

Qy    292 LKICSLKPKQPGHSKTQRPEEMPISNL 318
           |:| ||:|:|||||:||  ::: ::|
Db    327 ATRKSSRRSEP--NVQSKEEMTINIL 351
```

RESULT 8
```
P2YR_MELGA
ID   P2YR_MELGA         STANDARD;       PRT;   362 AA.
AC   P49652;
DT   01-FEB-1996 (Rel. 33, Created)
```

Best Local Similarity 29.1%; Pred. No. 3e-19;
Matches 95; Conservative 58; Mismatches 130; Indels 44; Gaps 7;

```
Qy    18  MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVVLFNLAVADFLMICLPFRTDYYLR   77
              :|:.|:||||:||:  |||: |:|:  ||||||::||||::|| ||||||:: ||:|
Db    43  LPTVYILVFITGFLGNSVAIMPVFHMRPWSGISVYMFNLALADFLYLTLPALIFYFN  102

Qy    78  RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC  137
              :.|||||:|:.: |  |||||:|| |||||||||:||:.:   :|| :| ||| |
Db   103  KTDWIFGDVWMCKLQRPIFHVNLYGSILFLTCISVHRYTGVVHPLKSLGRLKKNAVVYSS  162

Qy   138  TLWALVIL-----------GTVYLLLENHLCVQBTA-------VSCESFIMESANGWHD  178
              ||:|||:               |||:|:||:|:|:|          :|:  :|:
Db   163  LVWALVVAVIAPILFYSGTGVRRNKTITCYDTTADEVLRSYFVYSMCTTVFM------  214

Qy   179  IMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFMVVAIVFITCYLP---  235
              | :.:|.:|: |  |:: |||  ||:|:  :|:.:|::|:    :|:||||:||
Db   215  ------FCIPFIVILGCYGLVKALIYK-DLDNSPLRRKSIYLVIVLTVFAVSYLPFHV  267

Qy   236  ---SVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSSPPKFYNK  291
                 :.|: |||.||||  .:|   |||||:|||||:|| ::||:|::||:: ||||||::|::
Db   268  MKTLNLRARLDF-QTPQMCAFNDKVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR  326

Qy   292  LKICSLKPKQPGHSKTQRPEEMPISNL  318
              .  ||:|   :||  |:: :|
Db   327  ATRKSSRRSFP--NVQSKSEEMTLNIL  351
```

RESULT 9

P2YR_BOVIN

ID   P2YR_BOVIN   STANDARD;   PRT;   373 AA.
AC   P40042;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aortic endothelium;
RX   MEDLINE=95352058; PubMed=7626079;
RA   Henderson D.J., Elliot D.G., Smith G.M., Webb T.E., Dainty I.A.;
RT   "Cloning and characterisation of a bovine P2y receptor.";
RL   Biochem. Biophys. Res. Commun. 212:648-656(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus callosum;
RX   MEDLINE=99064562; PubMed=9848096;
RA   Deng G., Matute C., Kumar C.K., Fogarty D.J., Miledi R.;
RT   "Cloning and expression of a P2y purinoceptor from the adult bovine
RT   corpus callosum.";
RL   Neurobiol. Dis. 5:259-270(1998).
CC   -!- FUNCTION: Receptor for extracellular adenine nucleotides such as
CC   ATP and ADP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X87628; CAA60958.1; -.
DR   EMBL; U34841; AAC78275.1; -.
DR   PIR; JC4162; JC4162.

```
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     53     74       1 (POTENTIAL).
FT   DOMAIN       75     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    109       2 (POTENTIAL).
FT   DOMAIN      110    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    127    147       3 (POTENTIAL).
FT   DOMAIN      148    166       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    167    188       4 (POTENTIAL).
FT   DOMAIN      189    218       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    219    238       5 (POTENTIAL).
FT   DOMAIN      239    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    285       6 (POTENTIAL).
FT   DOMAIN      286    303       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    304    328       7 (POTENTIAL).
FT   DOMAIN      329    373       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    124    202       BY SIMILARITY.
FT   CARBOHYD     11     11       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC...)   (POTENTIAL).
SQ   SEQUENCE   373 AA;  42287 MW;  9270A7175C0BDA76 CRC64;
```

```
Query Match             19.6%;  Score 362.5;  DB 1;  Length 373;
Best Local Similarity   27.5%;  Pred. No. 3.9e-19;
Matches    92;  Conservative   68;  Mismatches  140;  Indels  35;  Gaps   7;

Qy     18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYLR 77
             |: :.|::.|||:|| ||    :|   |||||   :|||||||||||||||
Db     54 LPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADPLYVLTLPALIFYFN 113

Qy     78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLFVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
             ||||| ||| ||||  ||||| ||| ||||:|||:  :|||||||||| | ||| :
Db    114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLFCISAHRYSGVVYPLKSLGRLKKNAVYISV 173

Qy    138 TLWALVLGTVILLRNHLCVQET-AVSC----------ESFIMESANGWHDIMFQLEFFM 187
             .||:.||||.|| |:|||: .|: |:|            :|||::| |   | |: ||
Db    174 LVWLIVVGISPLFLFYSSGTGIRKNKTITCYDTTSDEVLRSYFIYSM------CTTVAMFCV 228

Qy    188 PLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP-------SVSAR 240
             || || ::|.:|:  |:|: |:| :|  ::  |:|:|||||| ||:|       ::||
Db    229 PLVLILGGCYGLIVRALLIYK-DLDNSPLRRKSIYLIVIILTVFAVSYIPFHVMKTMNLRAR 287

Qy    241 LYFLWTVPSSACDPSWFGALHITLSFTYMNSMLDPLVVYFSSPSFPPKFYNKLKICSLKPK 300
             ||:| ||  ::||:  || | ||| || | || | | |::  |:|  |
Db    288 LDF-QTPEMCAFNDRVTATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR---------- 337

Qy    301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSD 335
              :|||  :|||  |::|| | :::|: :|||: :|
Db    338 --ATRKASRRSEANLQSKSEDMTLNILSEFKQNGD 370
```

```
RESULT 10
CLT2_PIG
ID   CLT2_PIG        STANDARD;         PRT;   345 AA.
AC   Q95NO3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 2 (CysLTR2)
GN   CYSLTR2 OR CYSLT2.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA   Ohishi T., Soga T., Matsushina H., Furuichi K.;
RT   "Characterization of the cloned rat and porcine cysteinyl leukotriene
RT   receptors.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC       mediated via a G-protein that activates a phosphatidylinositol-
CC       calcium second messenger system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   --------------------------------------------------------------------------
DR   EMBL; AB052662; BAB60817.1; -.
DR   InterPro; IPR004071; Cysleuk_receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01533; CYSLTRECPTR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     43       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     44     64       1 (POTENTIAL).
FT   DOMAIN       65     73       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     74     94       2 (POTENTIAL).
FT   DOMAIN       95    124       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    125    145       3 (POTENTIAL).
FT   DOMAIN      146    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       4 (POTENTIAL).
FT   DOMAIN      176    205       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    206    226       5 (POTENTIAL).
FT   DOMAIN      227    246       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    247    267       6 (POTENTIAL).
FT   DOMAIN      268    287       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    288    308       7 (POTENTIAL).
FT   DOMAIN      309    345       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    112    188       BY SIMILARITY.
FT   CARBOHYD     20     20       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC...)   (POTENTIAL).
FT   CARBOHYD    184    184       N-LINKED (GLCNAC...)   (POTENTIAL).
SQ   SEQUENCE   345 AA;  39410 MW;  5D1B1FB89BB95905 CRC64;
```

```
Query Match             19.5%;  Score 362;  DB 1;  Length 345;
Best Local Similarity   31.3%;  Pred. No. 4e-19;
Matches    86;  Conservative   59;  Mismatches  118;  Indels  12;  Gaps   5;

Qy     19 PPLLIVAFVLGALGNGVALCGFCFHMKTWKPST---VYLFNLAVADFLLMICLPFRTDYY 75
             | :.|:|||:||||||||||   ||| :||    :|| ||||||||||||||||||||
Db     43 PIVLIVFIWGALGNGFSIIVF---LKPYKKSTSVNVPMLNLAISDLLFITLPFRVDYY 99

Qy     76 LRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVAADRYFKVVHPHHAVNTISTRVAAGI 135
             |:: ||  ||||||.|| || |:::||||||    :::|:|: |||.|:|:||||||||
Db    100 LRGSNXIFGDTPCRIMSYSMYVNMYSSIYFLTVLSVVRFLATVHFPRLLHCTSIK-NAWI 158

Qy    136 VCTLWALVLGTVILLRNHLCVQETAVSC---ESFIMESANGWHDIMFQLEFFMPLGII 192
             ||| |: |||::||| |||||||:|||    ||||||||||||||| | ||||| ||
Db    159 LCGVIWIFIMASSTVLLKNGSEQKDNVTLCLELNSNKVTKLKTMNYVALVVGFVLPFGTL 218

Qy    193 LFCSFKIVWSLRRRQQLARQARM--KKATRFMVVAIVFITCYLPSVSARLYFLWTVPSS 250
             || ||: ||| | || |  :::  :||| :|:||:|||:|| : | : :|: || |
Db    219 SICYLLIRALLKVEVPESGLRLSHRKALITVIIALIIFILLCFLPYHVLRTLHLLEMKAD 278

Qy    251 ACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSF 285
             : || ||| |  |||| || : |||||| : :|
Db    279 KCKDRLHKAVAVTLALAAANSCFNPLYFAGENF 313
```

```
             :| |:.: |||| ||: ||  :|.:||||:|||||||||| :: ||:||
Db     54 LPAVYIIVFIIGFLGNSVAIWMFVFHMKPWSGISVTMFNLALADFLYVLTLPALIFYFN 113

             :||| |. ||| || |.|.||:|.|||||||  ::|||.:||||:||:|||||| ||
Qy     78 RRHWAFGDIPCRVGLFPTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISCRVVAAGIVC 137

             . |||||: |::.  ||| :  ||:|:|| :|||:||::||||:|.: |:||:||:
Db    114 KTDWIFGDVWCKLQRPIFHVNLYGSLLFLTCISAHRYSGVVYPLKSLGRLKKKNAIYVSV 173

             |.|:||:|: ||||: ||.|:|:|:  :  :  ::       -----ESFIMESANGWHIDIMFQLSFFM
Qy    138 TLWALVILGTVYLLLENHLCVQST-AVSC------ESFIMESANGWHIDIMFQLSFFM 187

             |.||:||: |||.|  |:  |:|:|||:|||||:: ||:     ----CTTVAMFCI
Db    174 LVWLIVVAISPILFYSGTGRKNKVTCYDSTSDEYLRSYFIYSM----CTTVAMFCI 228

             || ||||:|:|| |. |:||||||||| |:| :|  |:|||| ||     -------SVSAR
Qy    188 PLGIILFCSFKIWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP-------SVSAR 240

             |:.|||. :|| |: .|:||| |.:| ::|:||:||:|:|||:||:|:||:.||
Db    229 PLVLLIGCYGLIVRALIYK-DLDNSPLRRKSIYLIVLIVLTVFAVSYIPFHVMKTMNLRAR 287

             ||:.  |  :: |  :  :|||.||: |::|||:  :| ||:|:|:|:::||
Qy    241 LYFLWTVPSSACD--PSVHGALHITLSFTYMNSMLDPLVYFSSPSPPKFYNKLKICSLK 298

             |:|      :.|: ::|||:.  :|||||:|:| |:|||:|||||.:|       ------
Db    288 LDP---QTPEMCDFNDRVIATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR------ 337

             |:|.| ||:|   |:|.:  :::|:|
Qy    299 PKQPGHSKTQRPEBMPISNLGRRSCISVANSFQSQSD 335

             ---ATRKASRRSRSEANLQSKSBEWTLNILSBFKQNGD 370
Db    338
```

RESULT 11
P2YR_RAT
ID   P2YR_RAT   STANDARD;   PRT;   373 AA.
AC   P49651;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Insulinoma;
RX   MEDLINE=95298025; PubMed=7779087;
RA   Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RT   "Cloning of rat and mouse P2Y purinoceptors.";
RL   Biochem. Biophys. Res. Commun. 211:211-218(1995).
CC   -!- FUNCTION: Receptor for extracellular adenine nucleotides such as
CC   ATP and ADP. In pancreatic islets, may mediate some of the effects
CC   of extracellular ATP on insulin secretion.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in muscle, heart, liver, kidney,
CC   lung, brain, spleen, but not in testis.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

DR   EMBL; U22830; AAA91303.1; -.
DR   HSSP; P34996; 1DDD.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN   1   52   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   53   74   1 (POTENTIAL).
FT   DOMAIN   75   87   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   88   109   2 (POTENTIAL).
FT   DOMAIN   110   126   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   127   147   3 (POTENTIAL).
FT   DOMAIN   148   166   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   167   188   4 (POTENTIAL).
FT   DOMAIN   189   218   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   219   238   5 (POTENTIAL).
FT   DOMAIN   239   265   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   266   285   6 (POTENTIAL).
FT   DOMAIN   286   303   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   304   328   7 (POTENTIAL).
FT   DOMAIN   329   373   CYTOPLASMIC (POTENTIAL).
FT   DISULFID   124   202   BY SIMILARITY.
FT   CARBOHYD   11   11   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   27   27   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   113   113   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   197   197   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   373 AA; 42321 MW; 6DDF67628?B5B648 CRC64;

Query Match   19.3%; Score 358.5; DB 1; Length 373;
Best Local Similarity   27.0%;
Matches   91; Conservative 68; Mismatches 139; Indels 39; Pred. No. 7.6e-19; Gaps 8;

Qy   18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADFLLMICLPFRTDYYLR 77
```

```
RESULT 12
P2YR_CAVPO
ID   P2YR_CAVPO   STANDARD;      PRT;   373 AA.
AC   P59902;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Hartley; TISSUE=Small intestine;
RA   Gao N., Hu H., Zhu M.X., Wood J.D.;
RT   "A novel P2Y1 receptor in the guinea pig submucous plexus.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for extracellular adenine nucleotides such as
CC       ATP and ADP. (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

DR   EMBL; AY048684; AAL05953.1; -.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     57       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     58     78       1 (POTENTIAL).
FT   DOMAIN       79     89       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     90    110       2 (POTENTIAL).
FT   DOMAIN      111    128       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    129    149       3 (POTENTIAL).
FT   DOMAIN      150    167       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    168    188       4 (POTENTIAL).
FT   DOMAIN      189    220       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    221    241       5 (POTENTIAL).
FT   DOMAIN      242    258       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    259    279       6 (POTENTIAL).
FT   DOMAIN      280    310       EXTRACELLULAR (POTENTIAL).
```

FT   TRANSMEM   311   331   7 (POTENTIAL).
FT   DOMAIN   332   373   CYTOPLASMIC (POTENTIAL).
FT   DISULFID   124   202   BY SIMILARITY.
FT   CARBOHYD   11   11   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   27   27   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   113   113   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   197   197   N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   373 AA;   42300 MW; 2BD66543FFE3F341 CRC64;

Query Match   19.2%;   Score 356.5; DB 1; Length 373;
Best Local Similarity   26.9%;   Pred. No. 1.1e-18;
Matches   90; Conservative   69; Mismatches 141; Indels   35; Gaps   7;

Qy   18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLR   77
               : :.|| .|| .|| | |||| ||:|.:: ||||||
Db   54 LPAVYIVWFIIGFLGNSIAIWMFVFHMKPWSGISVYMFNLALADFLYVTLPALIFYYFN 113

Qy   78 RRIHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVPHPHAVNTISTRVAAGIVC 137
               :.|| |||| :||: ||||||| | ::| |:|| | |||||:|| | |:
Db   114 KTNWIFGDAMCKLQREIFHVNLYGSILFLTCISAHRYSGVVYPLXSLGRLKKKNAVYISV 173

Qy   138 TLWALVILGTVYLLLENHLCVQET-AVSC------------BSFIMESANGWHDIMPQLEFFM 187
               :|.: |:: :: :::   :
Db   174 LVWLIVVVAISPILFEYSGTGIRKNKTITCYDTTSDEVLRSYFIYSM-----CTTVAMFCV 228

Qy   188 PLGIILFCSFKIVWSLRRRQOLARQARMKKATRPIMVVAIVFITCYLP-------SVSAR 240
               || ||||| |||| :| | : : | | ||
Db   229 PLVLILGCYGLIVRALIYK-DLDNSPLRRKSIYLVIIVILTVFAVSYIPPHVMKTMNLRAR 287

Qy   241 LYFLWLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYPSSPSFPKFYNKLKICSLKPK 300
               : | | | : :.| :||||||: | ||
Db   288 EDF-QTPEMCTFNDRVYATYQVTRGLASINSCVDPILYFLAGDTFRRRLSR---------- 337

Qy   301 QPGHSKTQRPEMPISNLGRRSCISVANSFQSQSD 335
               :|| | ::: |: |:
Db   338 --ATRKA.SRRSEANLQSKSEDMTLNILSEFKQWSD 370

RESULT 13
CLT2_RAT
ID   CLT2_RAT   STANDARD;   PRT;   309 AA.
AC   Q94T9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cysteinyl leukotriene receptor 2 (CysLTR2) (RSBPT32).
GN   CYSLTR2 OR CYSLT2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA   Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT   "Characterization of the cloned rat and porcine cysteinyl leukotriene
RT   receptors.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC   mediated via a G-protein that activates a phosphatidylinositol-
CC   calcium second messenger system (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AB052661; BAB60816.1; -.

DR   InterPro; IPR004071; Cysleuk receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01533; CYSLTRECPTR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN   1   26   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   27   47   1 (POTENTIAL)
FT   DOMAIN   48   56   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   57   77   2 (POTENTIAL)
FT   DOMAIN   78   98   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   99   119   3 (POTENTIAL)
FT   DOMAIN   120   137   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   138   158   4 (POTENTIAL)
FT   DOMAIN   159   187   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   188   208   5 (POTENTIAL)
FT   DOMAIN   209   229   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   230   250   6 (POTENTIAL)
FT   DOMAIN   251   271   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   272   292   7 (POTENTIAL)
FT   DOMAIN   293   309   CYTOPLASMIC (POTENTIAL).
FT   DISULFID   95   171   BY SIMILARITY.
FT   CARBOHYD   14   14   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   166   166   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   167   167   N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   309 AA;   35507 MW; A122AC8177879D56 CRC64;

Query Match   19.2%;   Score 355.5; DB 1; Length 309;
Best Local Similarity   31.2%;   Pred. No. 1e-18;
Matches   91; Conservative   57; Mismatches 131; Indels   13; Gaps   7;

Qy   2 YNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPST---VYLFNLA   58
               :: | : |||| |:|||||||||| |||.||
Db   10 YSDKNCTIE-NFKRDFYPIYLLIIFVWGALGNGFSIYVF---LQTYKKSTSVNVFMLNLA   65

Qy   59 VADFLLMICLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVV 118
               :| |:| |: |||| ||| || ||:||||||| ::| |:|||| | |:
Db   66 ISDFLFISTLPPRADYNFRGSDWIFGDWACRIMSYSLYVNMYTSIYFLTVLSIVRFLATA 125

Qy   119 HPHHAVNTISTRVAAGIVC-TLWALVILGTVYLLLENHLCVQETAVSCESFIMESAN--G 175
               || :||: | |: |:|| .|||.|||||||||||| |||:| : || ::||::|| |
Db   126 HPFQMLHITSVR-SAWILCGIIWVFIMASSGLLLKHGQEKKQNTTLCFELNLQKFKNIVI 184

Qy   176 WHDIMPQLEFFMPLGIILFCSFKIVWSLRRRQ--QLARQARMKKATRPIMVVAIVFITCY 233
               || : || : | :||: ||: :| :|:|| | | |||||||||||||||||| |
Db   185 LNYIALGVGFLLPFILTICYLLIIRVLLKVEIPESGPRIDAQRKALTTIVIAMIIFFLLCF 244

Qy   234 LPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYPSSPSF 285
               |||::||||| |||||| ||| ||||||||||||||||| ||||||| |
Db   245 LPYHALRTIHLVTWDADSCMDELHKATVITLTLAAANSCFNPFLYFAGENF 296

RESULT 14
P2YR_HUMAN
ID   P2YR_HUMAN   STANDARD;   PRT;   373 AA.
AC   P47900;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN   P2RY1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96257237; PubMed=8666290;
RA   Leon C., Vial C., Cazenave J.-P., Gachet C.;
RT   "Cloning and sequencing of a human cDNA encoding endothelial P2Y1

```
RT    purinoceptor.";
RL    Gene 171:295-297(1996).
RN    [2]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=96158962; PubMed=8579591;
RA    Ayyanathan K., Tania W., Harbansjit S., Raghbir A.S., Barnard E.A.,
RA    Kunapuli S.P.;
RT    "Cloning and chromosomal localization of the human P2Y1
RT    purinoceptor.";
RL    Biochem. Biophys. Res. Commun. 218:783-788(1996).
RN    [3]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=96205320; PubMed=8630005;
RA    Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,
RA    Boeynaems J.M.;
RT    "Cloning and tissue distribution of the human P2Y1 receptor.";
RL    Biochem. Biophys. Res. Commun. 221:588-593(1996).
RN    [4]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Blood;
RA    Leon C., Vial C., Weber J., Cazenave J.-P., Gacher C.;
RL    Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN    [5]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Lung;
RA    Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT    "cDNA clones of human proteins involved in signal transduction
RT    sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL    Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN    [6]
RP    SEQUENCE OF 95-373 FROM N.A., FUNCTION, AND INHIBITION STUDIES.
RC    TISSUE=Platelet;
RX    MEDLINE=98113162; PubMed=9442040;
RA    Jin J., Daniel J.L., Kunapuli S.P.;
RT    "Molecular basis for ADP-induced platelet activation. II. The P2Y1
RT    receptor mediates ADP-induced intracellular calcium mobilization and
RT    shape change in platelets.";
RL    J. Biol. Chem. 273:2030-2034(1998).
CC    -!- FUNCTION: Receptor for extracellular adenine nucleotides such as
CC        ATP and ADP. In platelets binding to ADP leads to mobilization of
CC        intracellular calcium ions via activation of phospholipase C, a
CC        change in platelet shape, and probably to platelet aggregation.
CC    -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC    -!- INDUCTION: Repressed by the P2Y1 receptor-specific antagonists
CC        A3P5P, A3P5P and A2P5P. These inhibit calcium ion mobilization
CC        and shape change in platelets.
CC    -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC    --------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------------
DR    EMBL; Z49205; CAA89066.1; -.
DR    EMBL; U42030; AAA97873.1; -.
DR    EMBL; U42029; AAA97872.1; -.
DR    EMBL; S81950; AAB47091.1; -.
DR    EMBL; AJ006945; CAA07339.1; -.
DR    EMBL; AY136752; AAN01278.1; -.
DR    EMBL; AF018284; AAB94556.1; -.
DR    PIR; JC4737; JC4737.
DR    HSSP; P34996; 1DDD.
DR    Genew; HGNC:8539; P2RY1.
DR    MIM; 601167; -.
DR    GO; GO:0005887; C:integral to plasma membrane; TAS.
DR    GO; GO:0004872; F:receptor activity; TAS.
DR    GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; TAS.
DR    GO; GO:0007200; P:G-protein signaling, coupled to IP3 second . . .; TAS.
DR    InterPro; IPR000276; GPCR_Rhodpsn.
DR    Pfam; PF00001; 7tm_1; 1.
DR    PRINTS; PR00237; GPCRRHODOPSN.
DR    PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR    PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW    G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;
KW    Blood coagulation.
FT    DOMAIN        1     52       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM     53     74       1 (POTENTIAL).
FT    DOMAIN       75     87       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM     88    109       2 (POTENTIAL).
FT    DOMAIN      110    126       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM    127    147       3 (POTENTIAL).
FT    DOMAIN      148    166       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM    167    188       4 (POTENTIAL).
FT    DOMAIN      189    218       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM    219    238       5 (POTENTIAL).
FT    DOMAIN      239    265       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM    266    285       6 (POTENTIAL).
FT    DOMAIN      286    303       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM    304    328       7 (POTENTIAL).
FT    DOMAIN      329    373       CYTOPLASMIC (POTENTIAL).
FT    DISULFID    124    202       BY SIMILARITY.
FT    CARBOHYD     11     11       N-LINKED (GLCNAC. . . .)   (POTENTIAL).
FT    CARBOHYD     27     27       N-LINKED (GLCNAC. . . .)   (POTENTIAL).
FT    CARBOHYD    113    113       N-LINKED (GLCNAC. . . .)   (POTENTIAL).
FT    CARBOHYD    197    197       N-LINKED (GLCNAC. . . .)   (POTENTIAL).
FT    CONFLICT    138    138       MISSING (IN REF. 1)
SQ    SEQUENCE   373 AA;  42071 MW;  4DC7C668B4145392 CRC64;
```

```
Query Match             19.1%;    Score 354;  DB 1;  Length 373;
Best Local Similarity   29.2%;    Pred. No. 1.6e-18;
Matches   93;  Conservative  66;  Mismatches 132;  Indels  28;  Gaps   8;

Qy      18  MPPLLIVAFVLGALGNGVALCGPCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYLR  77
                :|:|:|::|:|| |: | |      ::: |||:|||||| ::||  |:|
Db      54  LPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTLPALIFYFN  113

Qy      78  RRIHWAFGDIPCRVGLFTLAMNRAGSIVFLTVAADRYFKVHPHHAVNTISTRVAAGIVC  137
              |  |||       ::||:| |||| :|  |:  :  : |:
Db     114  KTDWIFGDAMCKLQRPIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKKKNAICISV  173

Qy     138  TLMALVILGTVYLLLENHLCVQET-AVSC------ESFIMESANGWHDIMFQL3FFM  187
               ::|::|     | :||:|:::|| :  ||:
Db     174  LVMLIVVVAISPILFYSGTGVRKNKTITCYDTTSDEYLRSYFIYSM-----CTTVAM?CV  228

Qy     188  PLGIILFCSFKIVWSLRRRQQLARQQARMKKATRFIMVVAIVFITCYLP------SVSAR  240
            || ||  | |:  :||:|  :|:||            :  |   |
Db     229  PLVILLGCYGLIVRALIYK-DLDNSPLRRKSIYLVIVLLTVFAVSYIPFHVMKTMNLRAR  287

Qy     241  LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYPSSPSPPKFYNKLKICSLKPK  300
             ||| |      |::||| | |  | |:      ::|||||  : :::|  :|:
Db     288  LDF-Q?PAMCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRR---RLSRATRKAS  343

Qy     301  QPGHSKTQ-RPEEMPISNL  318
            :|| |:| :|: |:
Db     344  RRSEANLQSKSEDMTLNIL  362
```

```
RESULT 15
P2YR_MOUSE
ID    P2YR_MOUSE              STANDARD;      PRT;   373 AA.
AC    P49650;
DT    01-FEB-1996 (Rel. 33, Created)
DT    01-FEB-1996 (Rel. 33, Last sequence update)
DT    28-FEB-2003 (Rel. 41, Last annotation update)
DE    P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN    P2RY1.
OS    Mus musculus (Mouse).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX    NCBI_TaxID=10090;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Insulinoma;
```

```
Db  231 VLILGCYGLIVKAL-IYNDLDNSPLRRKSIYLIVIVLTVFAVSYIPFHVMKTMNLRARLD 289
        :|||    ||  ||   |                    |      |      |  ||
Qy  243 FLWTVPSSACD--PSVHGALHITLSFTYMNSMLDPLVYFSSPSPPKFYNKLKICSLKPK 300

Db  290 F---QTPEMCDFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR-------- 337
        |        ||           :||     |        :|: ||  :|
Qy  301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSD 335

Db  338 --ATRKASRSEANLQSKSEEMTLNILSEFKQNGD 370
```

Search completed: June 30, 2004, 17:22:21
Job time : 18 secs

```
RX   MEDLINE=95298025; PubMed=7779087;
RA   Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RT   "Cloning of rat and mouse P2Y purinoceptors.";
RL   Biochem. Biophys. Res. Commun. 211:211-218(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RA   Leon C.;
RT   "Thromboresistance in P2Y1 receptor knockout mice.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for extracellular adenine nucleotides such as
CC       ATP and ADP. In pancreatic islets, may mediate some of the effects
CC       of extracellular ATP on insulin secretion.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U22829; AAA91302.1; -.
DR   EMBL; AJ245636; CAB57317.1; -.
DR   HSSP; P34996; 1DDD.
DR   MGD; MGI:105049; P2ry1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     52       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     53     74       1 (POTENTIAL).
FT   DOMAIN       75     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    109       2 (POTENTIAL).
FT   DOMAIN      110    126       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    127    147       3 (POTENTIAL).
FT   DOMAIN      148    166       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    167    188       4 (POTENTIAL).
FT   DOMAIN      189    218       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    219    238       5 (POTENTIAL).
FT   DOMAIN      239    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    285       6 (POTENTIAL).
FT   DOMAIN      286    303       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    304    328       7 (POTENTIAL).
FT   DOMAIN      329    373       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    124    202       BY SIMILARITY.
FT   CARBOHYD     11     11       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   373 AA;  42212 MW;  944125E9F4560BB3 CRC64;
```

Query Match 19.0%; Score 351.5; DB 1; Length 373;
Best Local Similarity 27.2%; Pred. No. 2.4e-18;
Matches 91; Conservative 66; Mismatches 143; Indels 35; Gaps 8;

```
Qy  18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVVLFNLAVADFLLMICLPFRTDYYLR 77
        :  | |:::|||  |  :||::|| |  ||| || :|||||:|||| ::|  |
Db  54 LPAVYILVLFIIGFLGNSVATWMPVFHMKPWSGISVVMFNLALADFLYVLTLPALIFYFN 113

Qy  78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHAVNTISTRVAAGIVC 137
        : :||||:|  |  | :| | |||:|:  || :|||| :|::   :|  |   ||:|
Db 114 KIDWIFGDAMCKLQRFIFHVNLYGSILFITCISAHRYSGVVYPLKSLGRLKKKNAIYVSV 173

Qy 138 TLWALVILGTVYLLLENHLCVQET-AVSCESFIMESANGWHDIMFQLE------FFMPL 189
        :|: ||  ::|: :   :   : :    :|                :|::|   |||
Db 174 LVWLIVVAISPILFYSGGTRKNKTVTCYD--TTSNDYLRSYFIYSMCTTVAMFCIPL 230

Qy 190 GIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP------SVSARLY 242
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:18:53 ; Search time 45 Seconds
(without alignments)
2425.984 Million cell updates/sec

Title: US-10-076-260-2
Perfect score: 1853
Sequence: 1 MXYNGSCCRIEGDTISQVMPP.........ANSFSQSDGQWDPHIVEWH 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	883.5	47.7	363	4	Q8TDS4	Q8tds4 homo sapien
2	881.5	47.6	387	4	Q8NGE4	Q8nge4 homo sapien
3	868.5	46.9	360	11	Q9EP66	Q9ep66 mus musculu
4	862.5	46.5	360	11	Q80239	Q80239 rattus norv
5	596	32.2	263	4	Q8NGV8	Q8ngv8 homo sapien
6	529	28.5	384	4	Q8NGW4	Q8ngw4 homo sapien
7	529	28.5	384	4	Q86WP7	Q86wp7 homo sapien
8	529	28.5	423	4	Q8TDS5	Q8tds5 homo sapien
9	450.5	24.3	319	4	Q9NQ20	Q9nq20 homo sapien
10	446.5	24.1	319	11	Q9JLS1	Q9jls1 mus musculu
11	375.5	20.3	309	11	Q8R528	Q8r528 mus musculu
12	351.5	19.0	390	13	Q8GGQ4	Q8ggq4 carassius a
13	348.5	18.8	373	11	Q8BMJ5	Q8bmj5 mus musculu
14	342.5	18.5	357	13	Q9DE05	Q9de05 raja erinac
15	341	18.4	361	13	Q90X57	Q90x57 xenopus lae
16	340.5	18.4	370	11	Q8BKK1	Q8bkk1 mus musculu
17	337	18.2	339	4	Q8N5S7	Q8n5s7 homo sapien
18	336.5	18.2	370	11	Q8BLG2	Q8blg2 mus musculu
19	336	18.1	390	13	Q8AXM7	Q8axm7 carassius a
20	330	18.1	347	13	Q7ZZA4	Q7zza4 brachydanio
21	329.5	17.8	380	13	Q9DGQ6	Q9dgg6 carassius a
22	325	17.8	309	4	Q86UH4	Q86uh4 homo sapien
23	325	17.5	344	4	Q7Z3S0	Q7z3s0 homo sapien
24	323.5	17.5	344	4	Q7Z3S6	Q7z3s6 homo sapien
25	317.5	17.5	377	13	Q7T2S9	Q7t2s9 carassius a
26	316	17.1	477	13	Q8JID5	Q8jid5 carassius a
27	309.5	17.1	307	11	Q8BS98	Q8bs98 mus musculu
28	308.5	16.7	307	11	Q8CB97	Q8cb97 mus musculu
29	307	16.6	337	4	Q8IV19	Q8iv19 homo sapien
30	303.5	16.6	359	6	Q9NOU1	Q9nou1 ovis aries
31	301	16.4	385	11	Q8BQ97	Q8bq97 mus musculu
32	300.5	16.2	346	6	Q95KS6	Q95ks6 ovis aries
33	300	16.2	367	6	Q8MI04	Q8mi04 ovis aries
34	298.5	16.2	385	4	Q86YF2	Q86yf2 homo sapien
35	298	16.1	375	11	Q8BYI1	Q8byi1 mus musculu
36	297	16.1	370	13	Q8UWL5	Q8uwl5 fugu rubrip
37	297	16.0	390	11	Q7TT86	Q7tt86 rattus norv
38	297	16.0	399	11	Q8R3I1	Q8r3i1 mus musculu
39	293	15.8	315	6	Q9GKP7	Q9gkp7 sus scrofa
40	293	15.8	374	13	O57466	O57466 meleagris g
41	293	15.8	385	11	Q9JK40	Q9jk40 mus musculu
42	292.5	15.8	361	6	O46685	O46685 bos taurus
43	292	15.8	317	11	Q99MT6	Q99mt6 mus musculu
44	292	15.8	359	6	Q9GLN9	Q9gln9 pan troglod
45	291	15.7	359	11	Q9EPP3	Q9epp3 cavia porce

ALIGNMENTS

RESULT 1
Q8TDS4
ID   Q8TDS4   PRELIMINARY;   PRT;   363 AA.
AC   Q8TDS4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative G-protein coupled receptor (Seven transmembrane helix
DE   receptor).
GN   GPCR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT   "Identification of G protein-coupled receptor genes from the human
RT   genome sequence.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA   Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT   "Genome-wide discovery and analysis of human seven transmembrane helix
RT   receptor genes.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22513958; PubMed=12522134;
RA   Wise A., Foord S.M., Fraser N.J., Barnes A.A., Elshourbagy N.,
RA   Eilert M., Ignar D.M., Murdock P.R., Steplewski K., Green A.,
RA   Brown A.J., Dowell S.J., Szekeres P.G., Hassall D.G., Marshall F.H.,
RA   Wilson S., Pike N.B.;
RT   "Molecular Identification of High and Low Affinity Receptors for
RT   Nicotinic Acid.";
RL   J. Biol. Chem. 278:9869-9874(2003).
RN   [4]
RP   SEQUENCE FROM N.A.

RESULT 2
Q8NGE4

```
ID   Q8NGE4     PRELIMINARY;     PRT;     387 AA.
AC   Q8NGE4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Seven transmembrane helix receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA   Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT   "Genome-wide discovery and analysis of human seven transmembrane helix
RT   receptor genes.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB065865; BAC06083.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   363 AA;   41849 MW;   C4B0EEC9CCB81D56 CRC64;

Query Match             47.7%;  Score 883.5;  DB 4;  Length 363;
Best Local Similarity   52.8%;  Pred. No. 4.6e-78;
Matches  180;  Conservative  47;  Mismatches  107;  Indels  7;  Gaps  4;

Qy      5  SCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLL  64
              :||| ::|:::|:.||::|| : : |.|:||||||||||||||||||
Db     17  NCCVFRDDFIVKVLPPVLGLEFIFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVADFLL  76

Qy     65  MICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAV  124
              :|||||: ||::||: |::||||||: :||:|.|||:||||:|||||||||:||||||||
Db     77  IICLPFLMDNYYVRRWDWKFGDIPCRLMLFMAMNRQGSIIFLTVVAVDRYFRVVHPHHAL  136

Qy    125  NTISTRVVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMFQLE  184
              |.||: ||||:||||:|| ::|||:|:|||::||   |:| |..:::.:.|||:|||| ||
Db    137  NKISNRTAAIISCLLWGITIGLTVHLLKKMPIQNGGANLCSSFSICHTFQWHEAMFLLE   196

Qy    185  FFMPLGIILFCSPKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSARLYFL  244
              || ||||||||:.:|:||||||:  :|: |:|:||||||||:||||:|:||||: ::|:|
Db    197  FFLPLGIILFCSARIIWSLRQR-QMDRHAKIKRAITFIMVAIVFVICFLPSVVVRIRIF   255

Qy    245  WTVPSSA---CD--PSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNLKICSLKP   299
              | :: |:    .  ||:::|:|||||||||||||||||::|||||:|:::||::::|:|
Db    256  WLLHTSGTQNCEVYRSVDLAPFITLSFTYMNSMLDPVVYYFSSPSFPNFFSTLINRCLQR  315

Qy    300  KQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDP  340
              |: |:.|.:|  ::.|.: ||::|||:|||:||:.:..:|.|
Db    316  KMTGEPDNNRSTSVELTGDPNKT-RGAPEALMANSGBPWSP  355
```

RESULT 3
Q9EP66

```
ID   Q9EP66     PRELIMINARY;     PRT;     360 AA.
AC   Q9EP66;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative seven transmembrane spanning receptor.
GN   PUMAG OR PUMA-G.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ, and C57BL/6;
RA   Schaub A., Futterer A., Pfeffer K.;
RT   "PUMA-G, an interferon-gamma inducible gene in macrophages is a novel
RT   member of the seven transmembrane spanning superfamily.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AJ300199; CAC17791.1; -.
DR   EMBL; AJ300198; CAC17790.1; -.
DR   MGD; MGI:1933383; Pumag.
DR   GO; GO:0016021; C:integral to membrane; IDA.
DR   GO; GO:0005525; F:GTP binding; IDA.
DR   GO; GO:0001614; F:purinergic nucleotide receptor activity; IDA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   360 AA;   41400 MW;   CCCE52A2475777FC CRC64;

Query Match             46.9%;  Score 868.5;  DB 11;  Length 360;
```

Previous result (continuation):

```
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor; Transmembrane.
SQ   SEQUENCE   387 AA;   44495 MW;   26433C855EB5EC81 CRC64;

Query Match             47.6%;  Score 881.5;  DB 4;  Length 387;
Best Local Similarity   52.2%;  Pred. No. 7.7e-78;
Matches  178;  Conservative  49;  Mismatches  107;  Indels  7;  Gaps  4;

Qy      5  SCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLL  64
              :||| ::|:::|:.||::|| : : |.|:||||||||||||||||||
Db     17  NCCVFRDDFIAKVLPPVLGLEFIFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVADFLL  76

Qy     65  MICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAV  124
              :|||||: ||::||: |::||||||: :||:|.|||:||||:|||||||||:||||||||
Db     77  IICLPFVMDYYVVRRSDWKFGDIPCRLVLFMFAMNRQGSIIFLTVVAVDRYFRVVHPHHAL  136

Qy    125  NTISTRVVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMFQLE  184
              |.||: ||||:||||:|| ::|||:|:|||::||   |:| |..:::.:.|||:|||| ||
Db    137  NKISNWTAAIISCLLWGITVGLTVHLLKKLLIQNGPANVCISFSICHTFRWHEAMFLLE   196

Qy    185  FFMPLGIILFCSPKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSARLYFL  244
              || ||||||||:.:|:||||||:  :|: |:|:||||||||:||||:|:||||: ::|:|
Db    197  FLLPLGIILFCSARIIWSLRQR-QMDRHAKIKRAITFIMVAIVFVICFLPSVVVRIRIF   255

Qy    245  WTVPSSA---CD--PSVHGALHITLSFTYMNSMLDPLVYFSSPSPFPKFYNLKICSLKP   299
              | :: |:    .  ||:::|:|||||||||||||||||::|||||:|:::||::::|:|
Db    256  WLLHTSGTQNCEVYRSVDLAPFITLSFTYMNSMLDPVVYYFSSPSFPNFFSTLINRCLQR  315

Qy    300  KQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDP  340
              |: |:.|.:|  ::.|.: ||::|||:|||:||:.:..:|.|
Db    316  KMTGEPDNNRSTSVELTGDPNKT-RGAPEALMANSGBPWSP  355
```

Also on page (RESULT 2 header reference block):

```
RA   Elshourbagy N.A.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB083632; BAB89345.1; -.
DR   EMBL; AB065876; BAC06094.1; -.
DR   EMBL; AY148884; AAN71621.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
```

```
Best Local Similarity   55.6%;   Pred. No. 1.3e-76;
Matches  178;  Conservative  39;  Mismatches  96;  Indels  7;  Gaps  4;

Qy    3 NG-SCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVAD   61
          : :||||| | |::||||:|| |||||| |||:|| ||||:|| :||||||||||||
Db   11 NGKNCCVFRDENIAKVLPPVLGLEFVFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVAD   70

Qy   62 FLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPH  121
          ||:|:||||| |||||| ||| |||||||||||:|||||||||:|||||||||:||||||
Db   71 FLLICLPFLTDNYVHNWDWRFGGIPCRVMFLMLAMNRQGSIIFLTVVAVDRYFRVVHPH   130

Qy  122 HAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMF  181
          |||| |||||||||| |||||||||| |||||:||||||||||||||||||||||||||
Db  131 HFLNKISNRTAAIISCFLWGLTIGLTVHLLYTNMMTKNGEAYLCSSFSICYNFRWHDAMF  190

Qy  182 QLEFFMPLGIILFCSFKIVWSLRRRQQLRQARQMKKATRFIMVVAIVFITCYLPSVSARL  241
          |||||||||||||:| : :||:||||| :|||:  ||||:|:: ||||||||||:::|:|
Db  191 LLEFFLPLAIILFCSGRIIWSLRQR-QMDRHAKIKRAINFIMVVAIVFIICFLPSVAVRI  249

Qy  242 YFLWTVPS--SACD--PSVHGALHITLSFTYMNSMLDPLVYYFSSPSPPKFYNKLKICS  296
          |||||||    || : ::::||:||||||||||||||||||||||: :|| |: |: :
Db  250 RIFWLLYKYNVRNCDIYSSVDLAFFTTLSFTYMNSMLDPVVYYFSSPSFPNFFSTCINRC 309

Qy  297 LKPKQPGHSKTQRPEEMPIS  316
          |:|::
Db  310 LRRKTLGEPDNNRSTSVELT  329
```

```
RESULT 4
Q80Z39
ID   Q80Z39;
AC   Q80Z39;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nicotinic acid receptor.
GN   HM74B.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Soga T., Kamhara M., Takasaki J., Matsumoto S., Saito T., Ohishi T.,
RA   Hiyama H., Matsuo A., Matsushime H., Furuichi K.;
RT   "Molecular identification of nicotinic acid receptor.";
RL   Biochem. Biophys. Res. Commun. 0:0-0(2003).
DR   EMBL; AB103062; BAC58009.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor; Transmembrane.
SQ   SEQUENCE   360 AA;  41458 MW;  975BDE8CA448A6C5 CRC64;
```

```
Query Match   46.5%;   Score 862.5;   DB 11;   Length 360;
Best Local Similarity   54.7%;   Pred. No. 5.2e-76;
Matches  175;  Conservative  43;  Mismatches  95;  Indels  7;  Gaps  4;

Qy    3 NG-SCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVAD   61
          : :||||| | |::||||:|| |||||| |||:|| ||||:|| :||||||||||||
Db   11 NGKNCCVFRDENIAKVLPPVLGLEFVFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVAD   70

Qy   62 FLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPH  121
          ||:|:||||| |||||| ||| |||||||||||:|||||||||:|||||||||:||||||
Db   71 FLLICLPFLTDNYVQNWDWRFGSIPCRVMFLMLAMNRQGSIIFLTVVAVDRYFRVVHPH   130

Qy  122 HAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMF  181
          |||| |||||||||| |||||||||| |||||:||||||||||||||||||||||||||
Db  131 HFLNKISNRTAAIISCFLWGITIGLTVHLLYTDMMTRNGDANLCSSFSICYTFRWHDAMF  190

Qy  182 QLEFFMPLGIILFCSFKIVWSLRRRQQLRQARQMKKATRFIMVVAIVFITCYLPSVSARL  241
          |||||||||||||:| : :||:||||| :|||:  ||||:|:: |||||||||:::|:|
Db  191 LLEFFLPLGIILFCSGRIIWSLRQR-QMDRHVKIKRAINFIMVVAIVFICFLPSVAVRI  249

Qy  242 YFLWTVPS--SACD--PSVHGALHITLSFTYMNSMLDPLVYYFSSPSPPKFYNKLKICS  296
          |||||||    || : ::::||:||||||||||||||||||||||: :|| |: |: :
Db  250 RIFWLLYKHVRNCDIYSSVDLAFFTTLSFTYMNSMLDPVVYYFSSPSFPNFFSTCINRC  309

Qy  297 LKPKQPGHSKTQRPEEMPIS  316
          |:|::
Db  310 LRRKTLGEPDNNRSTSVELT  329
```

```
RESULT 5
Q8NGV8
ID   Q8NGV8;
AC   Q8NGV8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Seven transmembrane helix receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA   Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT   "Genome-wide discovery and analysis of human seven transmembrane helix
RT   receptor genes.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB065665; BAC05891.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor; Transmembrane.
SQ   SEQUENCE   263 AA;  30289 MW;  F96BF023A5E1F1C1 CRC64;
```

```
Query Match   32.2%;   Score 596;   DB 4;   Length 263;
Best Local Similarity   59.9%;   Pred. No. 4.5e-50;
Matches  115;  Conservative  23;  Mismatches  54;  Indels  0;  Gaps  0;

Qy    5 SCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLL   64
          ||||||| |:|::||||:|: |||| |||:|| :|||||:|:| ::|::::||||||||||
Db   17 NCCVFRDDFIVKVLPPVLGLEPIFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVADFLL   76

Qy   65 MICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHAV  124
          :||||| :|:|  | ||: |||||:::|::|||||:||||||||||:|||||:||||||:
Db   77 IICLPFLMDNYVRRWDWKPGDIPCRLMLFMLAMNRQGSIIFLTVVAVDRYFRVVHPHAL  136

Qy  125 NTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMFQLE  184
          | |||||||: |||||||| ||||||||:  ||:  |: ||||||:| :::|::||:||||
Db  137 NKISNRTAAIISCLLWGITIGLTVHLLKKKMPIPDGGANLCSSFSICHTLQWHEAMFLLE  196

Qy  185 FFMPLGIILFCS  196
          ||:||||||||
Db  197 FFLPLGIILFCS  208
```

```
RESULT 6
Q8NGW4
ID   Q8NGW4;                           PRELIMINARY;        PRT;     384 AA.
AC   Q8NGW4;
```

RESULT 7
Q86WP7
ID   Q86WP7   PRELIMINARY;   PRT;   384 AA.
AC   Q86WP7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Seven transmembrane helix receptor.
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA   Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT   "Genome-wide discovery and analysis of human seven transmembrane helix
RT   receptor genes.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB065652; BAC05878.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; 7tm_1; 1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor; Transmembrane.
SQ   SEQUENCE   384 AA;   41426 MW;   1C845SFED8085F36 CRC64;

Query Match   28.5%; Score 529; DB 4; Pred. No. 2.4e-43;
Best Local Similarity   39.2%; Pred. No. 2.4e-43; -
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps   7;

Qy   4 GSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFL 63

Db   42 GPCHPTSSSLVSAFLAPILALEFVLGLVGNSLALFIFCIHTRPWTSNTVFLVSLVAADFL 101

Qy   64 LMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHA 123

Db   102 LISNLPLRVDYLLHETWRFGAAACKVNLFMLSTNRTASVVFLTAIALNRYLKVVQPHHV 161

Qy   124 VNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQE-TAVSCESFIM----ESANGWHD 178

Db   162 LSRASVGAAARVAGGLWVGI-----LLLNGHLLLSTFSGPSCLSYRVGTKPSASLRWHQ 215

Qy   179 IMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVAIVFITCYLPSV- 237

Db   216 ALYLLEFFLPLALILFAIVSIGLTIRNR-GLGGQAGPQRAMRVLAMVVAVYTICFLPSII 274

Qy   238 --SARLYFLWTVPSSA---CDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSF 285

Db   275 FGMASMVAFWLSACRSLDLCTQLFHG---SLAFTYLNSVLDPVLYCFSSPNF 323

RESULT 8
Q8TDS5
ID   Q8TDS5   PRELIMINARY;   PRT;   423 AA.
AC   Q8TDS5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative G-protein coupled receptor.
GN   GPCR OR TG1019.
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT   "Identification of G protein-coupled receptor genes from the human
RT   genome sequence.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22191290; PubMed=12065583;
RA   Hosoi T., Koguchi Y., Sugiwara E., Chikada A., Ogawa K., Tsuda N.,
RA   Suto N., Tsunoda S., Taniguchi T., Ohmuki T.;
RT   "Identification of a Novel Human Eicosanoid Receptor Coupled to
RT   Gi/o.";
RL   J. Biol. Chem. 277:31459-31465(2002).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB083630; BAB89343.1; -.
DR   EMBL; AB083055; BAC11806.2; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.

RT   and Neutrophils.";
RL   Mol. Pharmacol. 63:471-477(2003).
DR   EMBL; AY158687; AAO17739.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; 7tm_1; 1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   384 AA;   41412 MW;   0C5B35FED8085F36 CRC64;

Query Match   28.5%; Score 529; DB 4; Length 384;
Best Local Similarity   39.2%; Pred. No. 2.4e-43; -
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps   7;

Qy   4 GSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFL 63

Db   42 GPCHPTSSSLVSAFLAPILALEFVLGLVGNSLALFIFCIHTRPWTSNTVFLVSLVAADFL 101

Qy   64 LMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHA 123

Db   102 LISNLPLRVDYLLHETWRFGAAACKVNLFMLSTNRTASVVFLTAIALNRYLKVVQPHHV 161

Qy   124 VNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQE----ESANGWHD 178

Db   162 LSRASVGAAARVAGGLWVGI-----LLLNGHLLLSTFSGPSCLSYRVGTKPSASLRWHQ 215

Qy   179 IMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVAIVFITCYLPSV- 237

Db   216 ALYLLEFFLPLALILFAIVSIGLTIRNR-GLGGQAGPQRAMRVLAMVVAVYTICFLPSII 274

Qy   238 --SARLYFLWTVPSSA---CDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSF 285

Db   275 FGMASMVAFWLSACRSLDLCTQLFHG---SLAFTYLNSVLDPVLYCFSSPNF 323

```
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   423 AA;  45810 MW;  72ADA9F43A9EC051 CRC64;

Query Match             28.5%;  Score 529;  DB 4;  Length 423;
Best Local Similarity   39.2%;  Pred. No. 2.7e-43;
Matches  115;  Conservative  48;  Mismatches 108;  Indels  22;  Gaps    7;

Qy     4 GSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFL 63
             :|   :   | |  ||||||  || :|| :  ||  :||   :|| :| ||:|:|||||
Db    81 GPCHPTSSSLIVSAFLAPILALEFVLGLVGNSLALFIFCIHTRPWTSNTVFLVSLIVAADFL 140

Qy    64 LMICLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHA 123
            : ||      ||      || | |  :|:  :|| |:::||| | |:||::||:|  :
Db   141 LISSNLPLRVDYLLHETWRFGAAACKVNLFMLSTNRTASVFLTAIALNRYLKVVQPHEV 200

Qy   124 VNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQE-TAVSCESFIM---ESANGWHD 178
            :   |  ::  :  :|  || ||   ||| :   |:   |   | :    :|| ||:
Db   201 LSRASVGAAARVAGGLWVGI------LLLNGHLLLSTFSGPSCLSYRVGTKPSASLRWHQ 254

Qy   179 IMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSV- 237
            | :|||| ||  ||:  |   : |:|:| |  |  ::  : |:: ||| :  | |||:
Db   255 ALYLLEFFLPLALILFAIVSIGLTIRNR-GLGGQAGPQRAMRVLAMVVAVYTICFLPSII 313

Qy   238 --SARLYFLWTVPSSA---CDPSWHGALHITLSFTYMNSMLDPLVYYFSSPSF 285
             : | ||| ::  :    | |   : ::|  | ||:|| |||:|| |||| |
Db   314 FGMASMVAFWLSACRSLDLCTQLFHG----SLAFTYLNSVLDPVLYCFSSPNF 362
```

RESULT 9
Q9NQ20
```
ID   Q9NQ20      PRELIMINARY;      PRT;    319 AA.
AC   Q9NQ20;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   BA517H2.2 (G protein-coupled receptor 31) -
GN   GPR31.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Almeida J.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL121935; CAB99329.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   319 AA;  35074 MW;  2ACD0350AD7FB53A CRC64;

Query Match             24.3%;  Score 450.5;  DB 4;  Length 319;
Best Local Similarity   34.4%;  Pred. No. 9.6e-36;
Matches  101;  Conservative  61;  Mismatches 121;  Indels  11;  Gaps    5;

Qy     7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMI 66
            |   |  :  || :|:|    :| ::|||:||  : : :||| :|:||:| |:|||:||
Db     6 CSAPSTVVATAVGVLGLBCGLGLGNAVALWTFLFFVRVWKPYAVYLLNLALADLLLAA 65

Qy    67 CLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 126
            |||  : ::  :|| |   | : ||:  :  | :: : |  |: ||| :|||  :|:|
Db    66 CLPFLAAFYLSLQAWHLGRVGCWALHFLLDLSRSVGMAFLAAVALDRYLRVVHPRLKVNL 125

Qy   127 ISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANG-----WHDIMF 181
            :|   | |:   |  :   :    | | : ||:   :| :|:    : :      | :::
Db   126 LSPQAALGVSGLVWLLMVALTCPGLLISE--AAQNSTRCHSF-YSRADGSFSIIWQEALS 182

Qy   182 QLEFFMPLGIILFCSFKIVWSLRRR-QQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
            || |: |:|:|  |  :|  :|::| :|| | | : :|   |  |   |  |:|  :| |
Db   183 CLQFVLPFGLIVFCNAGIIRALQKRLREPEKQPKLQRAQALVTLVVVLFALCFLPCFLAR 242

Qy   241 --LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSSPFPKFYNKL 292
             | :    | : ||   :     |:| :|::| |:||: |  ||| :  ::::
Db   243 VLMHIFQNLGSCRALCAVAHTSDVTGSLTYLHSVLNPVVYCFSSPTFRSSYRRV 296
```

RESULT 10
Q9JLS1
```
ID   Q9JLS1      PRELIMINARY;      PRT;    319 AA.
AC   Q9JLS1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   G protein coupled receptor.
GN   TCP10C.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99431663; PubMed=10501965;
RA   Schimenti J.C.;
RT   "ORFless, intronless, and mutant transcription units in the mouse t
RT   complex responder (Tcr) locus.";
RL   Mamm. Genome 10:969-976(1999).
DR   EMBL; AF140708; AAF26668.1; -.
DR   MGD; MGI:98543; Tcp10c.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   319 AA;  35551 MW;  571F6DFB485BD7C4 CRC64;

Query Match             24.1%;  Score 446.5;  DB 11;  Length 319;
Best Local Similarity   32.5%;  Pred. No. 2.4e-35;
Matches  104;  Conservative  64;  Mismatches 133;  Indels  19;  Gaps    7;

Qy     7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMI 66
            |   |  :  || :|:|    :| ::|||:|   : : :||| :|:||:|:|:|||:||
Db     6 CSAASTVVETAVGTMTLECVLGLMGNAVALWTFFYRLKVWKPYAVYLFNILVADLLLAT 65

Qy    67 CLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 125
            | ||  : ::  :| |   | :  ::   :  | :: : |  |: ||| :|||  :|:|
Db    66 SVPFFAAFYLKGKTMKLGHMPCQLLLFLAFSCGVGVAFLMTVALDRYLHVVHPRLRVNL 125

Qy   127 ISTRVAAGIVCTLWALVILGTVYLLLENHLC-VQETAVSCBSPI---MESANGWHDIMF 181
            :|   | |:   |  :| :    | |  |   :::  :|  |:        | | :::|
Db   126 LSLRAAWGISSLIWLLMVLTPQNLL---TCRTTQNSTECPSFYPTGGTKAIATCQEVLF 182

Qy   182 QLEFFMPLGIILFCSFKIVWSLRRR-QQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
            || |: |:|::  |   :  || :| :|  | |: :   | |  | |:   ::|||: |
Db   183 FLQVLLPFGLISFCNSGLIRTLQKRLSESDKQPTIRRARVIVAIMLLFGLCFLPSVLTR 242

Qy   241 --LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLK 298
             | :    | : ||   :     |:| :|::| |:||: |  |||:    ::::|:|:||:
```

Db 243 VLVHIFQEFKSCSVQQAIMRASDIAGSLTCLHSTLSPAIYCFSNPAFTHSYRKV-LKSLR 301

Qy 299 PKQPGHSKTQRPEEMPISNL 318
            :: ||
Db 302 GR------RKAAESPSDNL 314

RESULT 11
Q8R528
ID   Q8R528;   PRELIMINARY;   PRT;   309 AA.
AC   Q8R528;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cysteinyl leukotriene 2 receptor.
GN   CYSLTR2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ishii S.;
RT   "Mouse CysLT2 Gene.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB058930; BAB86881.1; -.
DR   MGD; MGI:1917336; Cysltr2.
DR   GO; GO:0031631; F:cysteinyl leukotriene receptor activity; IDA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01533; CYSLTRECPTR.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   309 AA;   35226 MW; 96FACC6B8AF96974 CRC64;

Query Match   20.3%; Score 375.5; DB 11; Length 309;
  Best Local Similarity   32.7%; Pred. No. 2e-28;
  Matches   96; Conservative   56; Mismatches 125; Indels 17; Gaps   7;

Qy   2 YNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPST---VYLFNLA 58
          |||| ||: :|||||||| :: ::| ||| || ||:: : ::   ::|| |||
Db   10 YSNRRNCTIE-NFKKEFYPIIYLIIFFWGALGNGFSIYVF---LQTCKKSTSVNVFMLNLA 65

Qy   59 VADFILMICLPFRTEYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVV 118
          || |:| ||||| || || ::| ||| ||||||:::| ||||||| ::| ||||
Db   66 TSDFLFISTLPFRADYYFRGSNWIFGLLACRVMSYSLYVNMYTSIYFLITVLSVVRFLATV 125

Qy 119 HPHHAVNTISTRVAAGIVCTLWALVILGTVYILLLENHLCVQETAVSC-----ESFIMESA 173
          || ::| ::| |||||||| | | |||||:| ||| |||:||:| |||: |
Db 126 HPFRMEHVTSVR-SAWILCGIIWVFIMASSALLLVNGQEEKDNIISCLELSPQKF--KSL 182

Qy 174 NGWHDIMFQLBFMPLGIILPCSFKIVWSLRRRQ--QLARQARMKKATRFIMVVAIVFIT 231
          |||:|| :|:|:| ||| :| : ||||:||: ::|:|:| |||| | ||:| |
Db 183 LIMNHIAVAVGFLLPFLTNICYLLIIRILLKAEIPESGPRAAHRKALTTIVIAMITFLL 242

Qy 232 CYLPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFYFSSPSF 285
          ||| ::| |:| | || | || ||:|| : ||| |:: ||||:|:| : :|
Db 243 CFLPYHALRTLHLVTWDKDSCGDVLHKATVIFLTMAAANSCFNPFLYFAGENF 296

RESULT 22
Q8QGQ4
ID   Q8QGQ4;   PRELIMINARY;   PRT;   390 AA.
AC   Q8QGQ4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Type five-like somatostatin receptor.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lin X., Peter R.E.;
RT   "Cloning and expression of a type five-like somatostatin receptor in
RT   goldfish.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF252879; AAM18805.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00038; HLH_1; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   390 AA;   44136 MW; EC12119A4B6CF9A8 CRC64;

Query Match   19.0%; Score 351.5; DB 13; Length 390;
  Best Local Similarity   30.0%; Pred. No. 5.9e-26;
  Matches   97; Conservative   71; Mismatches 134; Indels 21; Gaps 10;

Qy   3 NGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADF 62
          |||| : | |: | |::| | :| :||: |||||||:||||
Db   30 NGSSMABEDST--KILAVIYLVVPVVGLTGNSLAIFVVLRYTKMKTATNMYILNLAVADE 87

Qy   63 LLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHH 122
          | | ||||: : || : : ||| | :| :   :::| ::|:| ||||:||||:
Db   88 LYILGLPFLTAHNM-LGYWPFGNFLCRILMWTDGSISQFTSFCLTVMSIDRYMAVVHPIR 146

Qy 123 AVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHD--IM 180
          : : | :| | ::|:| | ::| :: || | | :| | ||
Db 147 SARWRRPRVAKVINSMVWALSCLLTLPVII--YCDVQPELNTCNLSWPEPRDVWSTAFIL 204

Qy 181 F--QLEFFMPLGIILFCSFKIVWSLR--RRQQLARQARM-KKATRFIMVVAIVFITCYL 234
          :   : || | :: : | :| |:: : : ::| | | | :| | ::| |:|
Db 205 YTAMLGFFFPLMVICLCYLLIVIKVKSASARAGLSKRGRSEKKVTRMVVIIVVVFVLCWL 264

Qy 235 P-SVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFESSPSFPFKYNKL- 292
          | :: | | | | | | : ::| | || |: | |:| :   :| |:
Db 265 PFPFININLISTLPENSL---MTGIYFLTVILYVNSCANPLLYSFLSDNFKRSFQQVL 321

Qy 293 --KICSLKPKQPGHSKTQRPEE 312
           |:| |: | ||: |::
Db 322 CIHKVNGVSNGHPGREHLSRSQQ 344

RESULT 13
Q8BMJ5
ID   Q8BMJ5;   PRELIMINARY;   PRT;   373 AA.
AC   Q8BMJ5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   P2Y purinoceptor 1.
GN   P2RY1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Body;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium;
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;

```
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573 (2002).
DR   EMBL; AK030759; BAC27125.1; -.
DR   MGD; MGI:105049; P2ry1.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   373 AA;  42228 MW;  BA8812AB7847287C CRC64;

Query Match             18.8%;  Score 348.5;  DB 11;  Length 373;
Best Local Similarity   26.9%;  Pred. No. 1.1e-25;
Matches   90;  Conservative   67;  Mismatches  143;  Indels   35;  Gaps    8;

Qy          18 MPPLLIVAFVLGALGNGVALCGCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLR 77
Db          54 LPAVYILVFIIGFLGNSVAIWMYVFHMKPWSGISVYMFNLALADFLIYVLTIPALIFYFN 113

Qy          78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVYSV 137
Db         114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKKRNAIYYSV 173

Qy         138 TLWALVILGTVTVLLLENHLCVQ&T-AVSCESFIMESANGWHDIMFQLE------FFMPL 189
Db         174 LVWLIVVAISPILFYSGTGTRKNKTVTCYD--TTSNDYLRSYFIYSMCTTVAMFCIPL 230

Qy         190 GIILFCSFKIVWSLRRRQQLARQARMKXATRFIMVVAIVFITCYLP------SVSARLY 242
Db         231 VLILGCYGLIVKAL-IYNDLDNSPLRRKSIYLVIIVLIVTVFAVSYIPFHVMKTMNLRARiD 289

Qy         243 FLWTVPSSACD--PSWHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK 300
Db         290 F---QTPEMCDFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR-------- 337

Qy         301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSD 335
Db         338 --ATRKASRRSEANLQSKSEEMTLNILSEFKQNGD 370


RESULT 14
Q9DE05
ID   Q9DE05      PRELIMINARY;      PRT;   357 AA.
AC   Q9DE05;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   P2Y receptor.
OS   Raja erinacea (Little skate).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=7782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=20459151; PubMed=10900200;
RA   Dranoff J.A., O'Neill A.F., Franco A.M., Cai S.Y., Connolly G.C.,
RA   Ballatori N., Boyer J.L., Nathanson M.H.;
RT   "A primitive ATP receptor from the little skate Raja erinacea.";
RL   J. Biol. Chem. 275:30701-30706 (2000).
DR   EMBL; AF242850; AAG2684.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   357 AA;  41239 MW;  14604E815DCBDB41 CRC64;

Query Match             18.5%;  Score 342.5;  DB 13;  Length 357;
Best Local Similarity   26.4%;  Pred. No. 4.1e-25;
Matches   87;  Conservative   69;  Mismatches  144;  Indels   29;  Gaps    9;

Qy          18 MPPLLIVAFVLGALGNGVALCGCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLR 77
Db          38 LPIMYIIVFVTGFIGNSVALWMFIFHMRPWSSITIYMFNLVLADLFYVFSLPILIFYFN 97

Qy          78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
Db          98 KTDWIFGELLCKLXRFIFHVNLYGSILFLTCISVHRYTGVVHPMKSLGRLKKK-SATIVC 156

Qy         138 T-LWALVILGTVYLLL--------ENHLCVQETAVS-CESFIMESANGWHDIMFQLEFF 186
Db         157 VCVWITVMAGISPILYFSRTGLRRNKTNTCYDTTSKELLETYFIYSMSTTF------FGPC 211

Qy         187 MPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP-------SVSA 239
Db         212 IPFATILVCYGFIVKALISNDM--KTPLRGKSVRLVIILVFAISYLPFHVMKNLNLQS 269

Qy         240 RLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKL--KICSL 297
Db         270 RLYY-QGLDTCEWNRRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRFTNAASREMTR 328

Qy         298 KPKQPGHSKTQRPEEMPISNLGRRSCISV 326
Db         329 RSELSMQFRSEDSPLQPVSNISQNGDTSL 357


RESULT 15
Q90X57
ID   Q90X57      PRELIMINARY;      PRT;   361 AA.
AC   Q90X57;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   P2Y1 nucleotide receptor.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cheng A.W. Tsim K.W.;
RT   "Cloning of Xenopus P2Y1 Receptor.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF432354; AAL27614.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   361 AA;  41002 MW;  E5B2D605F5B57FBD CRC64;

Query Match             18.4%;  Score 341;  DB 13;  Length 361;
Best Local Similarity   29.0%;  Pred. No. 5.8e-25;
Matches   93;  Conservative   61;  Mismatches  135;  Indels   32;  Gaps    9;

Qy          18 MPPLLIVAFVLGALGNGVALCGCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLR 77
Db          42 LPAVYIVVCITGFIGNSVAIWMFIFHMKPWSSISVYMFNLALADFLYVLSLPALIFYFN 101
```

```
Qy    78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
            : :     ||::  :| :|||||: || : :||:||| : :: : :|
Db   102 KTDWIFGDALCKLQRFLFHVNLYGSILFLTCISVHRYTGVVHPLKSLGRLKKKNSIYISA 161

Qy   138 TLMALVILGTVYLLLENHLCVQBT-AVSC-------ESFIMESANGWHDIMFQLEFFM 187
            :|:::|    :|     ::| :|
Db   162 LVWFIVIAGISPILFFSGTGIRKNKTITCFDTSSDEYLRSYFIYSM-----CTTVFGFCI 216

Qy   188 PLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP--------SVSAR 240
            :|||::|:::| : :| :|:|||   |||
Db   217 PFILILGCYGLIVRALIYKDMNRAPLR-KKSIYLVIIVLTVFAVSYLPFHVMKNLNLRAR 275

Qy   241 LYFLWTVPSSAC--DPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNKLKICSLK 298
            |   :  :  :| :||:::::| :|: :  :
Db   276 LDF---QSPEMCNFNDRVVATYQVTRGLASLNSCVDPILYFLAGDTFRR--KLSRATRK 329

Qy   299 PKQPGHSKTQ-RPEEMPISNL 318
            : : : ||::: :|
Db   330 ASRRSEANVQSKSEVTLNIL 350
```

Search completed: June 30, 2004, 17:23:21
Job time : 48 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:26:22 ; Search time 23 Seconds
(without alignments)
776.634 Million cell updates/sec

Title: US-10-076-260-2
Perfect score: 1853
Sequence: 1 MYNGSCCRIBGDTISQVMPP..........ANSFQSQSDGQWDPHIVEWH 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	881.5	47.6	387	4	US-09-170-496D-222	Sequence 222, App
2	880.5	47.5	387	4	US-09-170-496D-108	Sequence 108, App
3	529	28.5	423	2	US-08-955-713-2	Sequence 2, Appli
4	510	27.5	476	2	US-08-955-713-4	Sequence 4, Appli
5	451.5	24.4	319	3	US-09-130-749-2	Sequence 2, Appli
6	451.5	24.4	319	3	US-09-130-749-2	Sequence 2, Appli
7	448.5	24.2	319	4	US-09-170-496D-60	Sequence 60, Appl
8	448.5	24.2	319	4	US-09-170-496D-196	Sequence 196, App
9	371	20.0	362	2	US-08-513-974B-374	Sequence 374, App
10	362.5	19.6	373	3	US-08-559-524A-4	Sequence 4, Appli
11	362.5	19.6	373	4	US-08-749-707-4	Sequence 4, Appli
12	362.5	19.6	373	4	US-09-947-922-4	Sequence 4, Appli
13	361.5	19.5	346	4	US-09-585-876-2	Sequence 2, Appli
14	346.5	18.7	370	3	US-08-781-250-2	Sequence 2, Appli
15	336	18.1	339	1	US-08-153-848-44	Sequence 44, Appl
16	336	18.1	339	3	US-08-812-871-3	Sequence 3, Appli
17	336	18.1	339	3	US-09-299-843A-44	Sequence 44, Appl
18	336	18.1	339	4	US-09-088-337B-44	Sequence 44, Appl
19	336	18.1	339	4	US-09-170-496D-32	Sequence 32, Appl
20	336	18.1	339	5	PCT-US93-11153-44	Sequence 44, Appl
21	336	18.1	339	5	PCT-US95-07180-2	Sequence 2, Appli
22	335	18.1	339	4	US-09-170-496D-182	Sequence 182, App
23	328.5	17.7	302	4	US-08-467-948A-30	Sequence 30, Appl
24	328.5	17.7	302	3	US-08-467-947A-30	Sequence 30, Appl
25	327	17.6	309	3	US-09-422-869-20	Sequence 20, Appl
26	325	17.5	344	2	US-08-467-948A-8	Sequence 8, Appli
27	325	17.5	344	3	US-08-467-947A-8	Sequence 8, Appli
28	315.5	17.0	374	4	US-09-102-710B-3	Sequence 3, Appli
29	314	16.9	373	3	US-08-513-974B-373	Sequence 373, App
30	308	16.6	325	2	US-08-118-270-51	Sequence 51, Appl
31	308	16.6	325	5	PCT-US93-08528-51	Sequence 51, Appl
32	307	16.6	391	1	US-07-816-283-4	Sequence 4, Appli
33	307	16.6	391	1	US-08-417-103-4	Sequence 4, Appli
34	307	16.6	395	1	US-08-097-938-5	Sequence 5, Appli
35	307	16.6	395	1	US-08-476-000-5	Sequence 5, Appli
36	307	16.6	395	1	US-08-472-840-5	Sequence 5, Appli
37	307	16.6	395	2	US-08-476-976-5	Sequence 5, Appli
38	307	16.6	395	3	US-08-474-410-5	Sequence 5, Appli
39	306	16.5	398	1	US-08-097-938-6	Sequence 6, Appli
40	306	16.5	398	1	US-08-476-000-6	Sequence 6, Appli
41	306	16.5	398	1	US-08-472-840-6	Sequence 6, Appli
42	306	16.5	398	2	US-08-476-976-6	Sequence 6, Appli
43	306	16.5	398	3	US-08-474-410-6	Sequence 6, Appli
44	306	16.5	398	3	US-08-486-673B-6	Sequence 6, Appli
45	305.5	16.5	369	3	US-08-120-601B-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-170-496D-222
; Sequence 222, Application US/09170496D
; Patent No. 655339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SEQ ID NO 222
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-222

Query Match 47.6%; Best Local Similarity 52.2%; Score 881.5; DB 4; Length 387;
Matches 178; Conservative 49; Mismatches 107; Indels 7; Gaps 4;

Qy   5 SCCRIBGDTISQVMPPLLIVAFVIGALGNGVALCGFCFIMKTWKPSTVYLFNLAVADFLL 64
              ||: |:|: :|| ||||:|| |:| :| |:||: |||||:|||||
Db   17 NCCVFRDDFIAKVLPPVLGLEFIFPGLLGNGLALWIFCFHLKSWKSSRIFIFNLAVADFLL 76

Qy   65 MICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAV 124
             :||||:| : ||::||:|||||:| |:||:|||||| :||||:|||||:||: :||
Db   77 IICLPFVMDYYVRRSDWKPGDIPCRLVLFMFAMNRQGSIIFLTVVAVDRYFRVVHPHHAL 136

Qy   125 NTISTRVAAGIVCTLWALVILGTVYLLLLENHLCVQETAVSCESPIMESANGWHDIMFQLE 184
             | |: |||:|| :|| |:|||| ||:||:|||||| : : |||:| :::||::|:||
Db   137 NKISNWTAAIISCLLWGITVGLTVHLLKKKLLIQNGPANVCISFSICHTFRWHEAMFLLE 196

Qy   185 FFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSARLYFL 244
             || ||||||| :|||::|:::| :|| || :||:||||| || | | | ||||||| :|
Db   197 FLLPLGIILFCSARIIWSLRQR-QMDRHAKIKRAKTFIMVVAIVFVICFLPSVVVRIRIF 255

Qy   245 WTVPSSA---CD--PSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKP 299
                   | || |: |:| : |||: |||| ||||:|||||||| :| |::|| ::||
Db   256 WLLHTSGTQNCEVYRSVDLAFFITLSFTYMNSMLDPVVYYFSSPSFPNFFSTLINRCLQR 315

Qy   300 KQPGHSKTQRPEMPISNLGRRSCISVANSFQSQSDGQWDP 340
              | | |::| : : ::|: :| | |: |:| | | |
Db   316 KMTGEPDNNRSTSVELTGDPNKT-RGAPEALMANSGRPWSP 355

RESULT 2
US-09-170-496D-108
; Sequence 108, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 655539-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 108
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-108

Query Match   47.5%;   Score 880.5; DB 4;   Length 387;
Best Local Similarity 52.2%; Pred. No. 4.4e-68;
Matches 178; Conservative 49; Mismatches 167; Indels 7; Gaps 4;

Qy   5 SCCRIBGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADFLL   64
Db   17 NCCVFRDDFIAKVLPPVLGLEFIFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVADFLL   76

Qy   65 MICLPFRTDYILRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAV 124
Db   77 IICLPFVMDYYVRRSDWNFGDIPCRLVLFMFAMNRQGSIIFLTVVAVDRYFRVVHPHHAL 136

Qy 125 NTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMFQLE 184
Db 137 NKISNWTAAIISCLLWGITVGLTVHLLKKKLLIQNGPANVCISFSICHTFRWHEAMFLLE 196

Qy 185 FFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSARLYFL 244
Db 197 FLLPLGIILFCSARIIWSLRQR-QMDRHAKIKRAITFIMVAIVFVICFLFSVVRIRIF   255

Qy 245 WTVPSSA---CD--PSVHGALHITLSFTYMNSMLDPLVYFSSPSPKFYNKLKICSLKP   299
Db 256 WLLHTSGTQNCEVYRSVDLARPFITLSFTYMNSMLDFVVYYFSSPSFPNFFSLINRCLQR 315

Qy 300 KQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDP 340
Db 316 KMTGEPDNNRSTSVELTGDPNKT-RGAPERLMANSGEPWSP 355

RESULT 3
US-08-955-713-2
; Sequence 2, Application US/08955713
; Patent No. 5955308
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
; APPLICANT: MOONEY, JEFFREY
; APPLICANT: BERGSMA, DERK
; APPLICANT: HALSEY, WENDY
; TITLE OF INVENTION: cDNA CLONE HEOAD54 THAT ENCODES   A HUMAN 7-TRANS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,713
; FILING DATE: 23-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,124
; FILING DATE: 18-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70087
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-955-713-2

Query Match   28.5%;   Score 529; DB 2;   Length 423;
Best Local Similarity 39.2%; Pred. No. 8.1e-38;
Matches 115; Conservative 48; Mismatches 108; Indels 22; Gaps 7;

Qy   4 GSCCRIBGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADFL   63
Db   81 GPCHZTSSSLVSAFLAPILALEFVLGLVGNSLALFIFCIHTRPWTSNTVFLVSLVAADFL 140

Qy   64 LMICLPFRTDYILRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHA 123
Db 141 LISNLPLRVDYILHETWRFGAAACKVNLFMLSTNRTASVVFLTAIALNRYLKVVQPHHV 200

Qy 124 VNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQE-TAVSCESFIM----ESANGWHD 178
Db 201 LSRASVGAAARVAGGLWVGI-----LLLNGHLLLSTFSGPSCLSYRVGTKPSASLRWHQ   254

Qy 179 IMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSV- 237
Db 255 ALVLLEFFLPLALILFAIVSIGLTIRNR-GLGGQAGPQRAMRVLAMVVAVYTICFLPSII 313

Qy 238 --SARLYFLWTVPSSA---CDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSF 285
Db 314 PGMASMVAFWLSACRSLDLCTQLFHG----SLAFTYLNSVLDPVLYCFSSPNF 362

RESULT 4
US-08-955-713-4
; Sequence 4, Application US/08955713
; Patent No. 5955308
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
; APPLICANT: MOONEY, JEFFREY
; APPLICANT: BERGSMA, DERK
; APPLICANT: HALSEY, WENDY
; TITLE OF INVENTION: cDNA CLONE HEOAD54 THAT ENCODES   A HUMAN 7-TRANS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,713
FILING DATE: 23-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/050,124
FILING DATE: 18-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70087
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-713-4

Query Match   27.5%; Score 510; DB 2; Length 476;
Best Local Similarity 39.6%; Pred. No. 4e-36;
Matches 110; Conservative 47; Mismatches 101; Indels 20; Gaps   6;

Qy   18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLMICLPFRTDYYLR 77
           | :|| :|||| ||:||| ||| | ||| :| :||| || :| ||||
Db   6 LAPILALEFVLGLVGNSLALFIFCIHTRPWTSNTVFLVSLVAADPLLISNLPLRVDYYLL 65

Qy   78 RRHWAFGDIPCRVGLPTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
          ||| |||||| | | |:||: |||||| ||||||||| | :|: || |||||||| |
Db   66 HETWRFGAAACKVNLFMLSTNRKASVVFLTAIALNRYLKVVXPHHVLNRASVGAXARVXG 125

Qy 138 TIWALVIIGTVYLLLENHLCVQETAVSCESPIM----ESANGWHDIMFQLEFFMPLGIIL 193
         || ::|:| ||| || :::| :| :|| :|||||||| |||||
Db 126 GIWGILLLNGXILLNTF-----SGPSCLSYRVGTKFSASLRWHQALYLLEFFLPLALLL 180

Qy 194 FCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSV---SARLYFLWTVPSS 250
         || || :||| || | | |||||||||||| ||:||||||||| | || ||||
Db 181 FAIVSIGLTIRNR-GLGGQAGPQRAMRVLAMVVAVYTICFLPSIIFGWASMVAFWLSACR 239

Qy 251 A--CDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSF 285
         :   ::: |:| |:| || :|| |||||||||
Db 240 SLDLCTQLFHG-----SLAFTYLNSVLDPVLYCFSSPKF 273

RESULT 5
US-09-130-749-2
; Sequence 2, Application US/09130749
; Patent No. 6031090
GENERAL INFORMATION:
APPLICANT: SHABON, USMAN
APPLICANT: ELSHOURBAGY, NABIL
TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM
RECEPTOR (GPR31A)
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/130,749
FILING DATE: 07-Aug-1998

CLASSIFICATION: UNKNOWN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-70513
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-130-749-2

Query Match   24.4%; Score 451.5; DB 3; Length 319;
Best Local Similarity 34.4%; Pred. No. 2.8e-31;
Matches 101; Conservative 61; Mismatches 121; Indels 11; Gaps   5;

Qy   7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLMI 66
          | | :|| | ||||||||| ||||||||||| | || || || ||||||||||||
Db   6 CSAPSTVWATAVGVLLGLECGLGLLGNAVALWTFLFRVRVWKPYAVVLLNLALADLLLAA 65

Qy   67 CLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 126
         ||| ||| :|:|||||| | | ||| |||||||||||||||||||| |: || |||
Db   66 CLPFLAAFYLSLQAWHIGRVGCWALRFLLDLSRSVGMAFLAAVALDRYLRVVHPRLKVNL 125

Qy 127 ISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANG----WHDIMF 181
         ||: :| | || ||| :| ||::|||||| || :::| :|:|| || ||
Db 126 LSPQAALGVSGLVMLLMVALTCPGLLISE--AAQNSTRCHSF-YSRADGSFSIIWQEALS 182

Qy 182 QLEFEMPLGIILFCSFKIVWSLRRR-QQLARQARMKKATRPIMVVAIVFITCYLPSVSAR 240
         ||| ||||| || || |:|||| ||||| ||| || | |||||| ||||||| ||
Db 183 CLQFVLPFGLIVFCNAGITRALQKRLREPEKQPKLQRAQALVTLVVLFALCFLPCFLAR 242

Qy 241 --LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPPKFYNKL 292
          : :: :| :::   ::: ||:| || :|| ||||||| | |||
Db 243 VLMHIFQNLGSCRAICAVAHTSDVTGSLTYLHSVLMPVVYCFSSPTFRSSYRRV 296

RESULT 6
US-09-130-749-2
; Sequence 2, Application US/09130749
; Patent No. 6031344
GENERAL INFORMATION:
APPLICANT: SHABON, USMAN
APPLICANT: ELSHOURBAGY, NABIL
TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM
TITLE OF INVENTION: RECEPTOR (GPR31A)
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/130,749
FILING DATE: 07-AUG-1998
CLASSIFICATION:

```
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: PRESTIA, PAUL F
;     REGISTRATION NUMBER: 23,031
;     REFERENCE/DOCKET NUMBER: GP-70513
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 610-407-0700
;     TELEFAX: 610-407-0701
;     TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 319 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-130-749-2

Query Match             24.4%;  Score 451.5; DB 3;  Length 319;
Best Local Similarity   34.4%;  Pred. No. 2.8e-31;
Matches  101;  Conservative   61; Mismatches 121;  Indels   11;  Gaps    5;

Qy      7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMI 66
            ::  ::::: |:|||:|||:| ||:||: |||:|| | :::| |||:||||:||:||||
Db      6 CSAPSTVVATAVGVLLGLECGLGLLGNAVALWTFLFRVRVWKPYAVYLLNLALADLLLAA 65

Qy     67 CLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 126
            |||| |::|: : |  ||::|| ||  |  ::  :|::|:||||||:||:|||| ||:||
Db     66 CLPFLAAFYLSLQAWHLGRVGCWALRFLLDLSRSVGMAFLAAVALDRYLRVVHPRLKVNL 125

Qy    127 ISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANG-----WHDIMF 181
          | :  | |:||:||:| ||| ||:|:|| | |:||:||:|:|::  :        |:::|
Db    126 LSPQAALGVSGLVWLLMVALTCPGLLISE--AAQNSTRCHSF-YSRADGSFSIIWQEALS 182

Qy    182 QLEFFMPLGIILFCSFKIVWSLRRR-QQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
          :|::| ::|:|:|:: :||| ::   ||:|||:|:| :  ||:|||::: | | ||:|:|
Db    183 CLQFVLPFGLIVFCNAGIIRALQKRLREPEKQPKLQRAQALVTLVVVLFALCFLPCFLAR 242

Qy    241 --LYFLWTVPSSACDPSVHGALHITILSFTYMNSMLDPLVYFSSPSPPKFYNKL 292
              ||:|| ||:  ||  | :||:|| |:|: : ||:||:||:: : :| | :|
Db    243 VLMHIFQNLGSCRALCAVAHTSDVTGSLTYLHSVVNPVVYCFSSPTFRSSYRRV 296

RESULT 7
US-09-170-496D-60
; Sequence 60, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
;   APPLICANT: Behan, Dominic P.
;   APPLICANT: Chalmers, Derek T.
;   APPLICANT: Liaw, Chen W.
;   TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein
;   TITLE OF INVENTION: Receptors
;   FILE REFERENCE: AREN-0040
;   CURRENT APPLICATION NUMBER: US/09/170,496D
;   CURRENT FILING DATE: 1998-10-13
;   NUMBER OF SEQ ID NOS: 294
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-170-496D-60

Query Match             24.2%;  Score 448.5; DB 4;  Length 319;
Best Local Similarity   34.0%;  Pred. No. 5e-31;
Matches  100;  Conservative   62; Mismatches 121;  Indels   11;  Gaps    5;

Qy      7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMI 66
            ::  ::::: |:|||:|||:| ||:||: |||:|| | :::| |||:||||:||:||||
Db      6 CSAPSTVVATAVGVLLGLECGLGLLGNAVALWTFLFRVRVWKPYAVYLLNLALADLLLAA 65
```

```
Qy     67 CLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 126
            |||| |::|: : |  ||::|| ||  |  ::  :|::|:||||||:||:|||| ||:||
Db     66 CLPFLAAFYLSLQAWHLGRVGCWALRFLLDLSRSVGMAFLAAVALDRYLRVVHPRLKVNL 125

Qy    127 ISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANG-----WHDIMF 181
          | :  | |:||:||:| ||| ||:|:|| | |:||:||:|:|::  :        |:::|
Db    126 LSPQAALGVSGLVWLLMVALTCPGLLISE--AAQNSTRCHSF-YSRADGSFSIIWQEALS 182

Qy    182 QLEFFMPLGIILFCSFKIVWSLRRR-QQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
          :|::| ::|:|:|:: :||| ::   ||:|||:|:| :  ||:|||::: | | ||:|:|
Db    183 CLQFVLPFGLIVFCNAGIIRALQKRLREPEKQPKLQRAQALVTLVVVLFALCFLPCFLAR 242

Qy    241 --LYFLWTVPSSACDPSVHGALHITILSFTYMNSMLDPLVYFSSPSPPKFYNKL 292
              ||:|| ||:  ||  | :||:|| |:|: : ||:||:||:: : :| | :|
Db    243 VLMHIFQNLGSCRALCAVAHTSDVTGSLTYLHSVVNPVVYCFSSPTFRSSYRRV 296

RESULT 8
US-09-170-496D-196
; Sequence 196, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
;   APPLICANT: Behan, Dominic P.
;   APPLICANT: Chalmers, Derek T.
;   APPLICANT: Liaw, Chen W.
;   TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protei
;   TITLE OF INVENTION: Receptors
;   FILE REFERENCE: AREN-0040
;   CURRENT APPLICATION NUMBER: US/09/170,496D
;   CURRENT FILING DATE: 1998-10-13
;   NUMBER OF SEQ ID NOS: 294
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 196
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-170-496D-196

Query Match             24.2%;  Score 448.5; DB 4;  Length 319;
Best Local Similarity   34.0%;  Pred. No. 5e-31;
Matches  100;  Conservative   62; Mismatches 121;  Indels   11;  Gaps    5;

Qy      7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMI 66
            ::  ::::: |:|||:|||:| ||:||: |||:|| | :::| |||:||||:||:||||
Db      6 CSAPSTVVATAVGVLLGLECGLGLLGNAVALWTFLFRVRVWKPYAVYLLNLALADLLLAA 65

Qy     67 CLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 126
            |||| |::|: : |  ||::|| ||  |  ::  :|::|:||||||:||:|||| ||:||
Db     66 CLPFLAAFYLSLQAWHLGRVGCWALRFLLDLSRSVGMAFLAAVALDRYLRVVHPRLKVNL 125

Qy    127 ISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANG-----WHDIMF 181
          | :  | |:||:||:| ||| ||:|:|| | |:||:||:|:|::  :        |:::|
Db    126 LSPQAALGVSGLVWLLMVALTCPGLLISE--AAQNSTRCHSF-YSRADGSFSIIWQEALS 182

Qy    182 QLEFFMPLGIILFCSFKIVWSLRRR-QQLARQARMKKATRFIMVVAIVFITCYLPSVSAR 240
          :|::| ::|:|:|:: :||| ::   ||:|||:|:| :  ||:|||::: | | ||:|:|
Db    183 CLQFVLPFGLIVFCNAGIIRALQKRLREPEKQPKLQRAQALVTLVVVLFALCFLPCFLAR 242

Qy    241 --LYFLWTVPSSACDPSVHGALHITILSFTYMNSMLDPLVYFSSPSPPKFYNKL 292
              ||:|| ||:  ||  | :||:|| |:|: : ||:||:||:: : :| | :|
Db    243 VLMHIFQNLGSCRALCAVAHTSDVTGSLTYLHSVVNPVVYCFSSPTFRSSYRRV 296

RESULT 9
US-08-513-974B-374
; Sequence 374, Application US/085l3974B
; Patent No. 6114139
; GENERAL INFORMATION:
;   APPLICANT: Hinuma, Shuji
;   APPLICANT: Hosoya, Masaki
;   APPLICANT: Fujii, Ryo
;   APPLICANT: Ohtaki, Tetsuya
```

```
Qy   18 MPPLLIVAFVLGALGNGVALCGPCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYLR 77
          :|:|:|:||::|:||| ||::|||: ||:: :|||::::|||::: ||
Db   43 LPTVYILVFITGPLGNSVAIWMFVFHMRPWSGISVTMFNLALADFLYVLTLPALIFYFN 102

Qy   78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYPKVVHPHHAVNTISTRVAAGIVC 137
          :||::|:| ::||||:::||::: :::|||||| |||: :: :|: || :: ||
Db  103 KTDWIFGDVMCKLQRFIFHVNLYGSILFITCISVHRYTGVVHPLKSLGRLKKKNAVYVSS 162

Qy  138 TLWALVILGTVLLENHL--------CVQBTA---------VSCRSFIMESANGWHD 178
          :||||: :: :|: :||
Db  163 LWWALVVAVIAPILFYSGIGVRRNKTITCYDTTADBYLRSYFVYSMCTTVFM------ 214

Qy  179 IMFQLFFMPLGIILFCSFKIVWSLRRRQOLARQARMKCATRFIMVVAIVFITCYLP--- 235
          :::::| :::| :: |:
Db  215 -----FCIPFIVILGCYGLIVKALIYK-DLDNSPLRRKSTYLIVILVLTVFAVSYLPFHV 267

Qy  236 ---SVSARLYFLMTVPSSACDPSVHGALHITLSFTYMNSMLDFLVYFSSPSPPKFYNK 291
              :: |||:||| |::|::| ||| ::::||
Db  268 MKTLNLRARLDF-QTPQMCAFNDKVYATYQVTRGLASLNSCVDPILYFLAGDTTFRRRLSR 326

Qy  292 LKICSLKPKQPGHSKTQRPBEMPISNL 318
          : :|||| :: |
Db  327 ATRKSSRRSEP--NVQSKSEEMTLNIL 351
```

RESULT 10
US-08-559-524A-4
; Sequence 4, Application US/08559524A
; Patent No. 5871963
; GENERAL INFORMATION:
;   APPLICANT: Conley, Pamela B.
;   APPLICANT: Jantzen, Hans-Michael
;   TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
;   STREET: 1800 M Street, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20036-5869
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/559,524A
;   FILING DATE: 15-NOV-1995
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Adler, Reid G.
;   REGISTRATION NUMBER: 30,988
;   REFERENCE/DOCKET NUMBER: 044481-5010-00-US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-467-7000
;   TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 373 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-559-524A-4

Query Match   19.6%; Score 362.5; DB 2; Length 373;
Best Local Similarity   27.5%; Pred. No. 1.5e-23;
Matches 92; Conservative 68; Mismatches 140; Indels 35; Gaps 7;

```
Qy   18 MPPLLIVAFVLGALGNGVALCGPCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYLR 77
          :|: :||: :|| |::||| :||| :::::||:::: |||||||||| :||::|||: ::
```

Query Match   20.0%; Score 372; DB 3; Length 362;
Best Local Similarity   29.1%; Pred. No. 2.6e-24;
Matches 95; Conservative 60; Mismatches 128; Indels 44; Gaps 7;

Db 54 LPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVILTLPALIFYYFN 113

Qy 78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137

Db 114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRTYGSVVYPLKSLGRLKKKNAVYISV 173

Qy 138 TLWALVILGTVYLLLENHLCVQET-AVSC--------ESFIMESANGWHDIMFQLEFFM 187

Db 174 LVWLIVVVGISPILFYSGTGIRKNKTITCYDTTSDEYLRSYFIYSM----CTTVAMFCV 228

Qy 188 PLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP------SVSAR 240

Db 229 PLVLILGCYGLIVRALIYK-DLDNSPLRRKSIYLVIVILTVFAVSYIPFHVMKTMNLRAR 287

Qy 241 LYFLMWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK 300

Db 288 LDF-QTPEMCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR-------- 337

Qy 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSD 335

Db 338 --ATRKASRRSEANLQSKSEDMTLNILSEFFKQNGD 370

RESULT 11
US-08-749-707-4
; Sequence 4, Application US/08749707
; Patent No. 6063582
; GENERAL INFORMATION:
;   APPLICANT: Conley, Pamela B.
;   APPLICANT: Jantzen, Hans-Michael
;   TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
;   STREET: 1800 M Street, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20036-5869
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/749,707
;   FILING DATE: 15-NOV-1996
;   CLASSIFICATION: 536
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Adler, Reid G.
;   REGISTRATION NUMBER: 30,988
;   REFERENCE/DOCKET NUMBER: 044481-5010-01-US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-467-7000
;   TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 373 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: <Unknown>
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-749-707-4

Query Match   19.6%; Score 362.5; DB 3; Length 373;
Best Local Similarity 27.5%; Pred. No. 1.5e-23;
Matches 92; Conservative 68; Mismatches 140; Indels 35; Gaps 7;

Qy 18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLR 77

Db 54 LPAVILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVILTLPALIFYYFN 113

Qy 78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137

(Right column)

Db 114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKKKNAVYISV 173

Qy 138 TLWALVILGTVYLLLENHLCVQET-AVSC--------ESFIMESANGWHDIMFQLEFFM 187

Db 174 LVWLIVVVGISPILFYSGTGIRKNKTITCYDTTSDEYLRSYFIYSM----CTTVAMFCV 228

Qy 188 PLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP------SVSAR 240

Db 229 PLVLILGCYGLIVRALIYK-DLDNSPLRRKSIYLVIVILTVFAVSYIPFHVMKTMNLRAR 287

Qy 241 LYFLMWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKPK 300

Db 288 LDF-QTPEMCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRLSR-------- 337

Qy 301 QPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSD 335

Db 338 --ATRKASRRSEANLQSKSEDMTLNILSEFFKQNGD 370

RESULT 12
US-09-947-922-4
; Sequence 4, Application US/09947922
; Patent No. 6680373
; GENERAL INFORMATION:
;   APPLICANT: Conley, Pamela B.
;   APPLICANT: Jantzen, Hans-Michael
;   TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
;   STREET: 1800 M Street, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20036-5869
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/947,922
;   FILING DATE: 07-Sep-2001
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/749,707
;   FILING DATE: 15-NOV-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Adler, Reid G.
;   REGISTRATION NUMBER: 30,988
;   REFERENCE/DOCKET NUMBER: 044481-5010-01-US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-467-7000
;   TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 373 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: <Unknown>
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-922-4

Query Match   19.6%; Score 362.5; DB 4; Length 373;
Best Local Similarity 27.5%; Pred. No. 1.5e-23;
Matches 92; Conservative 68; Mismatches 140; Indels 35; Gaps 7;

Qy 18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLR 77

Db 54 LPAVILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVILTLPALIFYYFN 113

US-08-781-250-2
; Sequence 2, Application US/08781250
; Patent No. 6010877
; GENERAL INFORMATION:
; APPLICANT: Sathe, Ganesh
; APPLICANT: Van Horn, Stephanie
; APPLICANT: Bergsma, Derk
; APPLICANT: Mao, Joyce Yue
; TITLE OF INVENTION: CDNA CLONE HE8CS41 THAT ENCODES A NOVEL 7-TRANSMEMBRAN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SmithKline Beecham
;   STREET: 709 Swedeland Road
;   CITY: King of Prussia
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19406
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/781,250
;   FILING DATE: 10-JAN-1997
;   CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: William T. Han,
;   REGISTRATION NUMBER: 34,344
;   REFERENCE/DOCKET NUMBER: ATG50043
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 610-270-5219
;   TELEFAX: 610-270-4060
;   TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 370 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-781-250-2

Query Match   18.7%;   Score 346.5; DB 3; Length 370;
Best Local Similarity   31.2%;   Pred. No. 3.5e-22;
Matches 95; Conservative 52; Mismatches 128; Indels 29; Gaps 11;

Qy   24 VAFVLGALGNGVAL-CGFCFHMKTWKPSTVYLPNLAVADFLLMICLPFRTDYYLRRRHWAF 83
             |||:||| ||:||| |||:|::::|||| || |::||:||||||:|||||:| ::|:|||
Db   48 VVFILGLITNSVSLFVFCFRMKMRSETAIFITNLAVSDLLFVCTLPFKI-FYNFNRHWPF 106

Qy   84 GDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCT-LWAL 142
             |:|||||:|||||||:|||||||| ||:||:||:|| |||||||||:|||||: ||:
Db   107 GDTLCKISGTAFLTNIYGSMLFLTCISVDRFLAIVYPFRS-RTIRTRRNSAIVCAGVWIL 165

Qy   143 VILGTVYLLENHLCVQBTAVSC-ESFIMESANGWHDIMFQLE-------FFMPLGIILF 194
             |: ||||||||| | |: ||||||||:|| :: |::| ||:|||||
Db   166 VLSGGISASLFSTNVNNATTCFGF---SKRVWKTYLSKITIPIEVGFIIPLILNVS 222

Qy   195 CSFKIVWSLRRRQQLAR-QARMKKATRFIMVVAIVFITCYLPSVSARLYFLWTVPSSAC- 252
             |||||| :||:||:|| |||:|:||||:||| ||||||||| ||||||||:|| |:||
Db   223 CSSVVLRTIRKPATLSQIGTNKKKVLKMTVHMAVFVVCFVPYNS--VLFLYALVRSQAI 280

Qy   253 -----DPSVHGALHITLSFTYMNSMLDPLVYFSSPSPPK-FYNKLKICSLKPKQPGHSK 306
                  |:: :|| |:|||| :|||||:||||||| ||| :||:| | :| |:|
Db   281 TNCFLERFAKIMYPITLCLATLNCCFDPFIYFTLESFQKSFYINAHI----RMESLFK 335

Qy   307 TQRP 310
          |||
Db   336 TETP 339

RESULT 13
US-09-585-876-2
; Sequence 2, Application US/09585876
; Patent No. 6586205
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 43239, A No. 6586205e1 GPCR-Like Molecule and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 5800-88
; CURRENT APPLICATION NUMBER: US/09/585,876
; CURRENT FILING DATE: 2000-06-01
; EARLIER APPLICATION NUMBER: 60/182,061
; EARLIER FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-585-876-2

Query Match   19.5%;   Score 361.5; DB 4; Length 346;
Best Local Similarity   29.9%;   Pred. No. 1.7e-23;
Matches 94; Conservative 67; Mismatches 136; Indels 17; Gaps 8;

Qy   3 NGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPST---VYLFNLAV 59
            ||| | | || :|||||||||||:: ||| ||:|| | : ||| : ||||:|
Db   3 NSRNCTIB-NFKREFPPIVYLIIFFWGVLGNGLSIYVF---LQPYKKSTSVNWFMLNLAI 82

Qy   60 ADFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVH 119
             ||||||:|||||||||||||| |||:|||||||||||||:|||||| :: | : |:||
Db   27 NSRNCTIB-NFKREFPIVYLIIFFWGVLGNGLSIYVF---LQPYKKSTSVNWFMLNLAI 142

Qy   83 SDLLFISTLPFRADYLRGSSWIFGDLACRIMSYSLVNMYSSIYFLTVLSVVRFLAMVH 142
             :||||| |:||| ||| ||:|||| |:||| | | || |:| | :|| |||
Db   143 PFRLLHVTSIR-SAWILCGIIWILIMASSIMLLDSGSEQNGSVTSCLELNLYKIAKLQTM 201

Qy   120 PHHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQBTAVSC---ESFIMESANGW 176
             |||||||||:|||||| ||||:|:||||||||||| | |: ||||||||:||
Db   143 HDIMFQLEFFMPLGIILFCSFKIVWSLRRRQ--QLARQARMKKATRFIMVVAIVFITCYL 234

Qy   177 QLARQARMKKATRFIMVVAIVFITCYL 201

Qy   202 NYIALVVGCLLPFFTLSICYLLIIRVLLKVEVPESGLRVSHRKALTIIITLLIFFLCFL 261

Qy   235 PSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKI 294

Qy   262 PYHTLRTVHLJTWKVGLCKDRLHKALVTLVLAAANACFNPILYYFAGENFK---DRLK- 317

Qy   295 CSLKPKQPGHSKTQ 308

Db   318 SALRKGHPQKAKTK 331

RESULT 14

RESULT 15
US-08-153-848-44
; Sequence 44, Application US/0815384B
; Patent No. 5759804
; GENERAL INFORMATION:
;   APPLICANT: Godiska, Ronald
;   APPLICANT: Gray, Patrick W.
;   APPLICANT: Schweikart, Vicki L.
;   TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
;   NUMBER OF SEQUENCES: 64
;   CORRESPONDENCE ADDRES:
;   ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;   ADDRESSEE: Bicknell
;   STREET: 6300 Sears Tower, 233 South Wacker Drive
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/153,848
;   FILING DATE:
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/977,452
;   FILING DATE: 17-NOV-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: No. 5759804and, Greta E.
;   REGISTRATION NUMBER: 35,302
;   REFERENCE/DOCKET NUMBER: 31794
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (312) 474-6300
;   TELEFAX: (312) 474-0448
;   TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 44:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 339 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-153-848-44

Query Match   18.1%; Score 336; DB 1; Length 339;
Best Local Similarity   30.5%; Pred. No. 2.5e-21;
Matches 97; Conservative 54; Mismatches 153; Indels 14; Gaps 9;

Qy   7 CRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTMKPSTVYLFNLAVADFLLMI 66
            | | : : |:| ||| :|:| :| ||:: :|||
Db   23 CGQETPLENMLPASFVLLDFILALVGNTLALWLFIRDHKSGTPANVFLMHLAVADLSCVL 82

Qy   67 CLPFRYDYILRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNT 126
            ||| |:| |||| ||: |||| |:| | |||: ::||| :||||
Db   83 VLPTRIVYHFSGNHWPFGEIACRLTGFLFYLNMYASIYFLTCISADRFLAIVHP---VKS 139

Qy   127 ISTR--VAAGIVCT-LWALVILGTVVLLENHLCVQETAVSCESFIMESANGWHDIMFQL 183
            ||: | | | : | || :: |:|
Db   140 LKLRRPLYAHLACAFLWVVVAVAMAPLLVSPGTVQTNETVVCLQLYREKASHHALVSLAV 199

Qy   184 EFFMPLGIILFCSFKIVWSLRRRQLARQARMK-KATRFIMVVAIVFITCYLP-SVSARL 241
            | ||| :||: : |: |:|| || |:|
Db   200 AFTFPP:TVTCYLLIIRSL--RQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSV 257

Qy   242 YFL-WTVPSSACDPSVHGAL--HITLSFTYMNSMLDPLVYFSSPSPKFYNKLKICSLK 298
            |:| | :| :| : |: |:: || :|:| |: |:
Db   258 YVLHYRSHGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKPRHALCNL-LCGKR 316

Qy   299 PKQPGHSKTQRPEEMPIS 316

Db   317 LKGPPPSFPGKTNESSLS 334
            |:| | :| :| : :|

Search completed: June 30, 2004, 17:24:28
Job time : 24 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    June 30, 2004, 17:19:42 ; Search time 21 Seconds
                                     (without alignments)
                                     1584.870 Million cell updates/sec

Title:           US-10-076-260-2
Perfect score:   1853
Sequence:        1 MYNGSCCRIEGTISQVMP2.........ANSFQSQSDGQWDPH7VBWH 346
Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:      283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR 78:*
                 1: Pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	880.5	47.5	387	2	I69202	G protein-coupled
2	364	19.6	362	2	S33733	G protein-coupled
3	362.5	19.6	373	2	JC4162	P2Y receptor - bov
4	354	19.1	373	2	JC4737	G protein-coupled
5	343.5	18.5	370	2	JC5549	heptahelical P2Y5-
6	339	18.3	308	2	I50241	G protein-coupled
7	324	17.5	344	2	T09508	intron 17 purinerg
8	320	17.3	373	2	A47556	ATP receptor P2u -
9	310.5	16.8	420	2	I51667	thrombin receptor
10	309.5	16.7	363	2	I57940	somatostatin recep
11	307	16.6	391	2	C41795	somatostatin recep
12	305.5	16.5	391	2	A41795	somatostatin recep
13	305.5	16.5	391	2	A39297	somatostatin recep
14	304.5	16.4	359	2	S15403	angiotensin II rec
15	302.5	16.3	369	2	B41795	somatostatin recep
16	302	16.3	364	2	JQ1488	bradykinin B2 rece
17	302	16.3	399	2	I48705	proteinase activat
18	300.5	16.2	369	2	A45291	somatostatin recep
19	300.5	16.2	369	2	D41795	somatostatin recep
20	300	16.2	328	2	I55450	G protein-coupled
21	300	16.2	384	2	A47249	brain-specific som
22	300	16.2	428	2	A44021	somatostatin recep
23	299.5	16.2	369	2	JC2083	somatostatin recep
24	299	16.1	388	2	JN0605	somatostatin recep
25	298.5	16.1	418	2	A46226	somatostatin recep
26	298.5	16.1	432	2	A43448	throbin receptor
27	298	16.1	428	2	S30508	probable G protein
28	297	16.0	397	2	S66518	proteinase-activat
29	296.5	16.0	346	2	S29248	somatostatin recep
30	296	16.0	359	2	A48857	angiotensin II rec
31	296	16.0	375	2	A54946	P-2U nucleotide re
32	295	15.9	384	2	JC4629	somatostatin recep
33	294	15.9	372	2	I38532	delta opioid recep
34	292.5	15.8	361	2	JC5653	G protein-coupled
35	292	15.8	371	2	JC5498	G protein-coupled
36	291.5	15.7	359	2	S44425	angiotensin II rec
37	291.5	15.7	359	2	JC2134	angiotensin II rec
38	290	15.7	372	2	S34592	delta opioid recep
39	289	15.6	359	2	JC1104	angiotensin II rec
40	288	15.5	355	2	A45177	chemokine (C-C) re
41	288	15.5	372	2	B48227	delta opioid recep
42	287.5	15.5	365	2	S68208	G protein-coupled
43	287.5	15.5	371	2	JC5796	probable chemoattr
44	286.5	15.5	359	2	JH0621	angiotensin II rec
45	286.5	15.5	423	2	JC7677	allatostatin recep

ALIGNMENTS

```
RESULT 1
I69202
G protein-coupled receptor HM74 - human
C;Species: Homo sapiens (man)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 19-May-2000
C;Accession: I69202
R;Nomura, H.; Nielsen, B.W.; Matsushima, K.
Int. Immunol. 5, 1239-1249, 1993
A;Title: Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte che
A;Reference number: I54751; MUID:94092629; PMID:7505609
A;Accession: I69202
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-387 <RES>
A;Cross-references: GB:D10923; NID:g219866; PIDN:BAA01721.1; PID:g219867
C;Genetics:
A;Gene: HM74
C;Superfamily: G protein-coupled receptor 4

Query Match             47.5%;  Score 880.5;  DB 2;   Length 387;
Best Local Similarity   52.2%;  Pred. No. 5.4e-74;
Matches  178;  Conservative  49;  Mismatches 107;  Indels 107;  Gaps  4;

Qy    5 SCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLL  64
          ||| |||| |:|||::| |||||||  |||||||||: |:|||||:||||||||||||
Db   17 NCCVFRDDFIAKVLPPVLGLEBFIGFGLLGNGLALWIFCFHLKSWKSSRIFLFNLAVADFLL  76

Qy   65 MICLPPRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAV 124
          ||||| |||::|||::|:|||||||: |:||||||||:|||||||::|||:||||||||
Db   77 IICLPFVMDYYVVRRSDWNFGDIPCRLVLFMFAMNRQGSIIFLTVVAVDRYFRVVHPHHAL 136

Qy  125 NTISTRVAAGIVCTLWALVILGTVYILLENHLCVQETAVSCESFIMESANGWHDIMFQLB 184
          | :|  :||:: ||||  |::|||:  ::|:|||::|:::|||||:| ||::|:|::||
Db  137 NKISNWTAAIISCLLWGTVGLTVGLIVHLLKKLLIQNGPANVCISFSICHTFRWHEAMFLLE 196

Qy  185 FFMPLGILLFCSFKIVWSLRRRQQLARQARMKKATRPIMVVAIVFITCYLPSVGSARLYFL 244
          |:|PLGILLFCS::IIWSLR:R Q::RHA:IK:AIT:IMVVAIVF:TC:LPSV    RI::
Db  197 FLLPLGILLFCSARIIWSLRQR-QMDRHAKIKRAITFIMVVAIVFVICFLPSVVVVRIRIF 255

Qy  245 WTVPSSA---CD--PSVHGALHITLSFTYMNSMLDPLVYFSSPSFPKFYNKLKICSLKP 299
          | |  :    |:  :::  |: ITLSFTYMNSMLDP:VY:FSSPS:P:F::  :  CL:
Db  256 WLLHTSGTQNCEVYRSVDLAFFITLSFTYMNSMLDPVVYYFSSPSPPNFFSTLJNRCLQR 315

Qy  300 KQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDP 340
          |:      ::  : ::|::       : ||::      : P
Db  316 KMTGEPDNNRSTSVELTGDPNKT-RGAPEALMANSGEPWSP 355

RESULT 2
S33733
G protein-coupled receptor - chicken
C;Species: Gallus gallus (chicken)
```

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
A;Accession: S33733
R;Webb, T.E.; Simon, J.; Krishek, B.J.; Bateson, A.N.; Smart, T.G.; King, B.F.; Burnstoc
FEBS Lett. 324, 219-225, 1993
A;Title: Cloning and functional expression of a brain G-protein-coupled ATP receptor.
A;Reference number: S33733; MUID:93285340; PMID:8508924
A;Accession: S33733
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-362 <WEB>
A;Cross-references: EMBL:X73268; NID:g395084; PIDN:CAA51716.1; PID:g395085
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match   19.6%;   Score 364;   DB 2;   Length 362;
Best Local Similarity   29.1%;   Pred. No. 4.7e-26;
Matches 95;   Conservative 58;   Mismatches 130;   Indels 44;   Gaps   7;

Qy   18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLFPRTDYYLR 77
               :| :|:|| | ||||:||:|:|| | :||:||| |||||||||
Db   18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLFPRTDYYLR 77

Qy   43 LPTVYILVFIITGFLGNSVAIWMFVFHMRPWSGISVYMFNLALADFLYVLTLPALIFYFN 102
               :|:|:|||:|||:||||||:| |:|:| |:|| |:||||||||||:|:|||:|||||
Db   43 LPTVYILVFIITGFLGNSVAIWMFVFHMRPWSGISVYMFNLALADFLYVLTLPALIFYFN 102

Qy   78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137

Qy   103 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISVHRYTGVVHPLKSLGRLKKKNAVYVSS 162
               |||||||:|:||||||||||||||||||||:|||| ||||:|||||||||||||||:|:
Db   114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKKKNAVYISV 173

Qy   138 TLWALVIL----------GTVYLLLENHLCVQETA--------------VSCESFIMESANGWED 178
               ||||||| |||| ||||||:|||:| :: :||||||| :||
Db   174 LVWLLVVGISPILFYSGTGVRRNKTITCYDTTSDEYLRSYFYYSM------CTTVAMFCV 228

Qy   163 LVWALVVAVIAPILFYSGTGVRRNKTITCYDTTADEVLRSYFYYSMCTTVFM------- 214
               |||||:|:|:||||||||||||||||||||||||::|||||||||||:|:|
Db   229 PLVLILGCYGLIVRALIYK-DLDNSPLRRKSIYLVIVLTVFAVSYIPFHVMKTMNLRAR 287

Qy   179 IMFQLEFMPLGIILZCSFKIWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP--- 235
               ::|:| |||||:: :|||| :|:|||||||||||||||||||||:||||
Db   188 PLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP--------SVSAR 240

Qy   215 ---FCIPFIVILGCYGLIVKALIYK-DLDNSPLRRKSIYLVIIVILTVFAVSYLPFHV 267
               |||| | ||| :| :||||||||||||||||||||||| ||::|| |||||| :|||
Db   241 LYFLMTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNK 300

Qy   236 ----SVSARLYFLMTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNK 291
               ||::|||:|||||||:||| ||:|||:| |||:|||| :|||| : |||:||
Db   288 LDF-QTPEMCAFNDRVYATYQVTRGLASLNSCVDPILYFLLAGDTFRRRLSR------ 337

Qy   268 MKTLNLRARLDF-QTPEMCAFNDKVYATYQVTRGLASLNSCVDPILYFLLAGDTFRRRLSR 326
               |: ::| |||| |||||||||||:|||||||||||||||||||:|||||||||||||||||
Db   301 QPGHSKTQRPEEMPISNLGRRSCISVANSPSQSD 335

Qy   292 LKICSLKPKQPGHSKTQRPEEMPISNL 318
               | ||:::|:||||||||||||||||| |
Db   338 --ATRKASRRSEANLQSKSEDMTLNLSEFKQNGD 370

Qy   327 ATRKSSRRSEP--NVQSKSEEWTLNIL 351
               ||||:|||| : |||| ::|||:|:

RESULT 3
JC4162
P2Y receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 12-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
C;Accession: JC4162
R;Henderson, D.J.; Elliot, D.G.; Smith, G.M.; Webb, T.E.; Dainty, I.A.
Biochem. Biophys. Res. Commun. 212, 648-656, 1995
A;Title: Cloning and characterisation of a bovine P2y receptor.
A;Reference number: JC4162; MUID:95352058; PMID:7626079
A;Accession: JC4162
A;Molecule type: mRNA
A;Residues: 1-373 <HEN>
A;Cross-references: EMBL:X87628; NID:g1032484; PIDN:CAA60958.1; PID:g1032485
A;Experimental source: aortic endothelial cell
C;Genetics:
A;Gene: bovp2y
C;Superfamily: ATP receptor P2u
C;Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein
F;52-77/Domain: transmembrane #status predicted <TM1>
F;88-111/Domain: transmembrane #status predicted <TM2>
F;124-150/Domain: transmembrane #status predicted <TM3>
F;171-191/Domain: transmembrane #status predicted <TM4>
F;214-237/Domain: transmembrane #status predicted <TM5>
F;261-282/Domain: transmembrane #status predicted <TM6>
F;305-328/Domain: transmembrane #status predicted <TM7>
F;11,27,113,197/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;258/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match   19.6%;   Score 362.5;   DB 2;   Length 373;

Best Local Similarity   27.5%;   Pred. No. 6.7e-26;
Matches 92;   Conservative 68;   Mismatches 140;   Indels 35;   Gaps   7;

Qy   18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLFPRTDYYLR 77
               :| :|:|| | ||||:||:|:|| | :||:||| |||||||||
Db   54 LPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVTMENLALADFLYVLTLPALIFYFN 113

Qy   78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKKKNAVYISV 173

Qy   138 TLWALVILGTVYLLLENHLCVQET-AVSC----------ESFIMESANGWHD:MFQLEFFM 187
               |||||||| |||||||||:|||:| |:   ::|:||||:|||||
Db   174 LVWLLVVGISPILFYSGTGIRKNKTITCYDTTSDEYLRSYFIYSM----CTTVAMFCV 228

Qy   188 PLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLP-------SVSAR 240
               |||||||||||||||||||||||||||||||||||||||||||||||| :::|:
Db   229 PLVLILGCYGLIVRALIYK-DLDNSPLRRKSIYLVIVLTVFAVSYIPFHVVKTMNLRAR 287

Qy   241 LYFLMTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYFSSSPSFPKFYNKLKICSLKPK 300
               ||||:|||:|||| ||:|||:| |||:|||| :|||| : |||:|| ||| ::
Db   288 LDF-QTPEMCAFNDRVYATYQVTRGLASLNSCVDPILYFLLAGDTFRRRLSR------ 337

Qy   301 QPGHSKTQRPEEMPISNLGRRSCISVANSPSQSD 335
               |||||||||||||||||||||||||||||||||| |
Db   338 --ATRKASRRSEANLQSKSEDMTLNLSEFKQNGD 370

RESULT 4
JC4737
G protein-coupled receptor P2Y1 - human
N;Alternate names: P2Y1 purinergic receptor; P2Y1 purinoceptor
C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence revision 16-Aug-1996 #text_change 17-Nov-2000
C;Accession: JC4737; JC4615; S54253
R;Janssens, R.; Communi, D.; Pirotton, S.; Samson, M.; Parmentier, M.; Boeynaems, J.M.
Biochem. Biophys. Res. Commun. 221, 588-593, 1996
A;Title: Cloning and tissue distribution of the human P2Y1 receptor.
A;Reference number: JC4737; MUID:96205320; PMID:8630005
A;Accession: JC4737
A;Molecule type: DNA
A;Residues: 1-373 <JAN>
A;Cross-references: GB:S81950; NID:g1839438; PIDN:AAB47091.1; PID:g1839439
R;Ayyanathan, K.; Webbs, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.; Kunapuli, S.
Biochem. Biophys. Res. Commun. 218, 783-788, 1996
A;Title: Cloning and chromosomal localization of the human P2Y1 purinoceptor.
A;Reference number: JC4615; MUID:96158962; PMID:8579591
A;Accession: JC4615
A;Molecule type: mRNA
A;Residues: 1-373 <AYY>
A;Cross-references: GB:U42029; NID:g1147730; PIDN:AAA97872.1; PID:g1147731
A;Experimental source: erythro leukemia cells
R;Leon, C.; Vial, C.; Cazenave, J.; Gachet, C.
submitted to the EMBL Data Library, May 1995
A;Description: Cloning of a human putative P2Y receptor.
A;Reference number: S54253
A;Accession: S54253
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-137,139-373 <LEO>
A;Cross-references: EMBL:Z49205; NID:g798835; PIDN:CAA89066.1; PID:g798836
C;Comment: This receptor belongs to a family of G protein-coupled receptors. It respon
C;Genetics:
A;Gene: p2Y1; GDB:P2RY1
A;Cross-references: GDB:677125; OMIM:601167
A;Map position: 3pter-3qter
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane pr
F;52-77/Domain: transmembrane #status predicted <TM1>
F;88-111/Domain: transmembrane #status predicted <TM2>
F;124-152/Domain: transmembrane #status predicted <TM3>
F;171-191/Domain: transmembrane #status predicted <TM4>
F;214-237/Domain: transmembrane #status predicted <TM5>

```
F;261-282/Domain: transmembrane #status predicted <TM6>
F;305-328/Domain: transmembrane #status predicted <TM7>
F;11,27,113,197/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;258,336/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predict
F;330,339/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predict
F;343/Binding site: phosphate (Ser) (covalent) (by protein kinase C and calmodulin-depen

Query Match          19.1%;  Score 354;  DB 2;  Length 373;
Best Local Similarity 29.2%;  Pred. No. 4.2e-25;
Matches   93;  Conservative   66;  Mismatches  132;  Indels   28;  Gaps    8;

Qy    18 MPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADFLLMICLPFRTDYYLR 77
         :  |||| ::: ::  ||| |||:: | |||  ||:|||||:|||||::| ||| ||:|
Db    54 LPAVYILVFIIGFLGNSVAIWMFVFHMKPWSGISVYMFNLALADFLYVLTLPALIFYYFN 113

Qy    78 RRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC 137
         : || |||  |: ::|  |:||  ||:|||.|:| ||  || | :   : :|  |
Db   114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRLKKKNAICISV 173

Qy   138 TLWALVILGTVYLLENHLCVQET-AVSC------BSFIMESANGWHDIMFQLEFFM 187
          :| |:::  : :   |  ::.                 :::  |   : || :
Db   174 LVWLIVVVAISPILFYSGTGVRKNKTITCYDTTSDEYLRSYFIYSM----CTTVAMFCV 228

Qy   188 PLGILFCSFKIVWSLRRRQQLARQQARMKKATRFIMVVAIVFITCYLP------SVSAR 240
         || :| |  ::  | :                    :  ::  .        ||   :
Db   229 PLVLLGCYGLIVARALIYK-DLDNSPLRRKSIYLVIIVLTVFAVSYIPFHVMKTMNLRAR 287

Qy   241 LYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPK 300
         | |      | :   :          ::    :||::|:::  : ::
Db   288 LDF-QTPAMCAFNDRYAYTYQVTRGLASLNSCVDPILYFLAGDTFRR---RLSRATRKAS 343

Qy   301 QPGHSKTQ-RPEEMPISNL 318
                :  : :
Db   344 RRSEANLQSKSEDMTINIL 362


RESULT 5
JC5549
heptahelical P2Y5-like receptor - human
C;Species: Homo sapiens (man)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 24-Sep-1999
C;Accession: JC5549
R;Janssens, R.; Boeynaems, J.M.; Godart, M.; Communi, D.
Biochem. Biophys. Res. Commun. 236, 106-112, 1997
A;Title: Cloning of a human heptahelical receptor closely related to the P2Y5 receptor.
A;Reference number: JC5549; MUID:97366605; PMID:9223435
A;Accession: JC5549
A;Molecule type: DNA
A;Residues: 1-370 <JAN>
A;Cross-references: DDBJ:AF005419; NID:g2240034; PIDN:AAB66322.1; PID:g2240035
C;Superfamily: ATP receptor P2u

Query Match          18.5%;  Score 343.5;  DB 2;  Length 370;
Best Local Similarity 31.0%;  Pred. No. 3.9e-24;
Matches   94;  Conservative   52;  Mismatches  130;  Indels   27;  Gaps   10;

Qy    24 VAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADFLLMICLPFRTDYYLRRRHWAF 83
         | |::|| ::.|:| | | |||: :   : ::||||:|||:: |||:  ||: : ||.|
Db    48 VVFIIGLITNSVSLFVFCFRMKRSETAIFITNLAVSDLLFVCTLPFKI-FYNFNRHWPF 106

Qy    84 GDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCT-LWAL 142
         || |: .  |  ::: .||::|||.|: || :||:|   :   :  || | |: :  |
Db   107 GDTLCKISGTAFLTNIYGSMLFLTCISVDRFLAIVYPFRS-RTIRTRRNSAIVCAGVWIL 165

Qy   143 VILGTVYLLENHLCVQETAVSCESFIMESANGWHDIMFQLE------PFMPLGIILFC 195
         |: |  : |          :||:    .          : :        :||| ::.|
Db   166 VLSGGISASLFSTTNVNNATTTC--FEGLSKRVWKTYLSKITIFIEVGFIIPLILNVSC 223

Qy   196 SFKIVWSLRRRQQLAR-QARMKKATRFIMVAIVFITCYLPSVSARLYFLWTVPSSAC- 252
         |  ::  |    :   ||   || :     ::|:   ::|  : :| :|    : .|
Db   224 SSVVLRTLRXPATLSQIGTNKKVLKMITVHMAVFVVCFVPYNS-VLFLYALVRSQAIT 281

Qy   253 ---DPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPK-FYNKLKICSLKPKQPGHSKT 307
            :    :     :  :   :|| :|::::   : ::
Db   282 NCFLERFAKIMYPITLCLATLNCCFDPFIYFFLESFQKSFYINAHI-----RMBSLFKT 336

Qy   308 QRP 310
         :
Db   337 ETP 339


RESULT 6
I50241
G protein-coupled receptor 6H1 - chicken
N;Alternate names: purinoceptor 6H1
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Jun-2000
C;Accession: I50241; JC4618
R;Kaplan, M.H.; Smith, D.I.; Sundick, R.S.
J. Immunol. 151, 628-636, 1993
A;Title: Identification of a G protein coupled receptor induced in activated T cells.
A;Reference number: I50241; MUID:93329058; PMID:8393036
A;Accession: I50241
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-308 <KAP>
A;Cross-references: GB:L06109; NID:g304383; PIDN:AAB06587.1; PID:g304384
R;Webb, T.E.; Kaplan, M.G.; Barnard, E.A.
Biochem. Biophys. Res. Commun. 219, 105-110, 1996
A;Title: Identification of 6H1 as a P2Y purinoceptor: P2Y5.
A;Reference number: JC4618; MUID:96190677; PMID:8619790
A;Accession: JC4618
A;Molecule type: mRNA
A;Residues: 1-308 <WEB>
A;Cross-references: GB:L06109; NID:g304383; PIDN:AAB06587.1; PID:g304384
A;Experimental source: T-cells
C;Comment: This receptor plays a role in T-cell activation.
A;Genetics:
A;Gene: P2Y5
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein
F;15-40/Domain: transmembrane #status predicted <TM1>
F;51-74/Domain: transmembrane #status predicted <TM2>
F;89-109/Domain: transmembrane #status predicted <TM3>
F;133-153/Domain: transmembrane #status predicted <TM4>
F;177-201/Domain: transmembrane #status predicted <TM5>
F;227-248/Domain: transmembrane #status predicted <TM6>
F;269-292/Domain: transmembrane #status predicted <TM7>

Query Match          18.3%;  Score 339;  DB 2;  Length 308;
Best Local Similarity 30.3%;  Pred. No. 8.3e-24;
Matches   91;  Conservative   56;  Mismatches  123;  Indels   30;  Gaps   10;

Qy    24 VAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADFLLMICLPFRTDYYLRRRHWAF 83
         : |||| ::.| |:   |  |::     ::||||:|||||: ||||:  |||:: |||
Db    23 MWFVLGLIANCVAIYIFFTLKVRNETTYMLNLAISDLLFVTLPFRI-YYFVVRNWPF 81

Qy    84 GDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVC-TLWAL 142
         ||: |: .. |  ::: .||:|||.|: ||| :|:| |    : :| ||| ||| | |:
Db    82 GDVLCKISVTLFYTNMYGSILFLTCISVDRFLAIVHPFRS-KTLRTKRNARIVCAVWIT 140

Qy   143 VILGTV---YLLENHLCVQETAVSCESFIMESANGWHDIMFQL------EFFMPLGII 192
         |: |:       : |: : : |  |::   :   |   : |:        |||||||:
Db   141 VLAGSTPASFFQSTNRQNNTEQRTCFENF--PESTWKTYLSRIVIFIEIVGFFIPLLN 197

Qy   193 LFCSFKIVWSLRRRQQLAR-QARMKKATRFIMVVAIVFITCYLP-SVSARLYFLW--TV 247
         : || : :  |    ::   ||   || :::::: ::::  ::|    : | ::
Db   198 VTCSTMVLRTLNKPLTLSRNKLSKKKVLKMIFVHLVIFCFCFVPYNITLLLYSLMRTQTW 257

Qy   248 PSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPKQPGHSKT 307
          ::   ||   ::::  |      :|||:|| : | :
Db   258 INCSVVTAVRTMYPVTLCIAVSNCCFDPIVYIFTSDTNSE-------LDKKQQVHQNT 308


RESULT 7
T09508
```

intron 17 purinergic receptor P2Y5 - human
N;Alternate names: G-protein coupled receptor
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C;Accession: T09508
R;Bohm, S.K.; Trumpp, A.; Khitin, L.M.; Kong, W.; Payan, D.G.; Bunnett, N.W.
submitted to the EMBL Data Library, April 1997
A;Description: The human purinergic receptor P2Y5 is encoded in intron 17 of the retinob
A;Reference number: Z16705
A;Accession: T09508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-344 <BOH>
A;Cross-references: EMBL:AF000546; NID:g2232068; PID:g2232069
C;Genetics:
A;Map position: 13
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match   17.5%; Score 324; DB 2; Length 344;
Best Local Similarity   30.8%; Pred. No. 2.3e-22;
Matches   88; Conservative   56; Mismatches 120; Indels   22; Gaps   10;

Qy   24 VAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYILRRRHWAF 83
                 || | ::: | |:| ||||||| |::| : ::|
Db   26 MVFVLGLWSNCVAIYIFICVLKVRNETTYMINLAMSDLLFVFTLPFRI-EYFTTRNWPF 84

Qy   84 GDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHAVNTISTRVAAGIVCT-LWAL 142
                 |:| | ||||: |:|| :: ||||| || :| : : |:| : | | |:
Db   85 GDLLCKISVMLFYTNMYGSILFLTCISVDRFLAIWYFKS-KTLRTKRNAKIVCTGVWLf 143

Qy   143 VILGT--VYLLLENHLCVQETAVSCESFIMESANGWHDIMFQL-------EFFMPLGILL 193
                  : ||:| | : | | | : | | |:|||| |
Db   144 VIGGSAPAVFVQSTHSQGNNASRACFENFPEAT--WKTYLSRIVIFIBIVGFFIPLIINV 201

Qy   194 FCSFKIVWSLRRRQQLARQARMK-KATRFIMVVAIVFITCYLP-SVSARLYFL--WTVP 248
                  | |||| : :: | | | | | : :|| | : || : |: | | |||
Db   202 TCSSMVLKTLTKPVTLSRSKINKTKVLKMIFVHLIIFCFCFVPYNINLILYSLVRTQTFV 261

Qy   249 SSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSSPFKFYNKLKI 294
                  | :| | | ||||||||||| | |:| | :
Db   262 NCSVVAAVRTMYPITLCIAVSNCCFDPIVYYFTSDT---IQNSIKM 304

RESULT 8
A47556
ATP receptor P2u - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C;Accession: A47556
R;Lustig, K.D.; Shiau, A.K.; Brake, A.J.; Julius, D.
Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993
A;Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.
A;Reference number: A47556; MUID:93281707; PMID:7685114
A;Accession: A47556
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-373 <LUS>
A;Cross-references: GB:L14751; NID:g309457; PIDN:AAA39871.1; PID:g309458
C;Superfamily: ATP receptor P2u
C;Keywords: transmembrane protein

Query Match   17.3%; Score 320; DB 2; Length 373;
Best Local Similarity   30.1%; Pred. No. 6e-22;
Matches   98; Conservative   44; Mismatches 150; Indels   34; Gaps   8;

Qy   4 GSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFL 63
                 :   :: : ||||||||| | |:: | | |:| :|||||||| | :|
Db   4 GYKCRFNEDFKYVLLPVSYGVVCVGLCLNVVALVIFLCRLKIWNASTTYMFHLAVSDSL 81

Qy   22 LMICLPFRTDYILRRRHWAFGDIPCRVGLFTLAMNRAGSIVF-ZVVAADRYFKVVHPHHA 123
                 |:|||||| :| :::::| ||:|| | | | | |: | : | : || |||| | ||:
Db   82 YAASLPLLVYYARGDHWPFSTVLCKLVRFLFYTNLYCSTLFLTCISVHRCLGVLRPLHS 141

Qy   124 VNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCES--------PIMESAN 174
                 : ||: || | : | | | | ||:: :|::: |:| :|:| |
Db   142 LRWGRARYARRVAAVVNVIVLACQAPVLYFVTTSVRGTRITCHDTSARELFSHFVAYSS- 200

Qy   175 GWHDIMFQLEFMPLGIILFCSFKIVWSLRRRQQ---------------ARQARMKKATRFIMVA 226
                 || | | | | :| ::|||:||| :::   :|| | | :|
Db   201 ----VMLGLLFAVPFSVILVC--YVLMARRLLKPAYGTTGGLPRAKR--KSVRTIALVL 251

Qy   227 IVFIICYLP-SVSARLYFLWTVPSSACDP--SVHGALHITLSFTYMNSMLDPLVYYFSSP 283
                 |:|:| :| : | ::|| : : :| |: : || |||||||||||||||| ||
Db   252 AVFALCFLPFHVTRTLYYSERSLDLSCHTLNAINMAYKITRPLASANSCLDPVLYFLAGQ 311

Qy   284 SFPKFYNKLKICSLKPKQPGHSKTQR 309
                   |||||||| |:||| | |
Db   312 RLVRFARDAK----PPTEPTPSPQAR 333

RESULT 9
I51667
thrombin receptor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51667
R;Gerszten, R.E.; Chen, J.; Ishii, M.; Ishii, K.; Nanevicz, T.; Turck, C.W.; Vu, T.H.;
Nature 368, 648-651, 1994
A;Title: Thrombin receptor's specificity for agonist peptide is determined by its extr
A;Reference number: I51667; MUID:94195429; PMID:8145852
A;Accession: I51667
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-420 <GER>
A;Cross-references: EMBL:U09632; NID:g9495197; PIDN:AAA18498.1; PID:g9435198

Query Match   16.8%; Score 310.5; DB 2; Length 420;
Best Local Similarity   28.0%; Pred. No. 5.2e-21;
Matches   88; Conservative   56; Mismatches 133; Indels   37; Gaps   10;

Qy   14 ISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTD 73
                 :: || | | ::: :|: | : | |: | ||||| | :| |: | |
Db   100 LTKFVPSLYTVVFIVGLPLNLLAIIIFLFKMKVRKPAVVMNLAIADVFFVSVLPFKIA 159

Qy   74 YYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHAVNTISTRVAA 133
                 || :|:: :|: |: | ||| ||:: |::| : : ||| : |:: | |: : ||
Db   160 YHLSGNDWLFGPGMCRIVTAIFYCNMYCSVLLIASISVDRFLAVVYPMHSLSW-R7MSRA 218

Qy   134 GIVCT-LWALVILGTVYLLL--ENHLCVQETAVSCBSFI-MESANGWHDIMFQ-----LEF 185
                 |:| | | ::||::| |: :   : | | | | | : |
Db   219 YMACSFIWLISIASTIPLLVTBQTQKIPRLDITTCHDVLDLKDFYIYFSSFCLLFF 278

Qy   186 FMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITC------YL 234
                 | || :|:| | :   : ||: || | | |: :|: |
Db   279 FVPFIITTICYIGIIRSL-SSSIENSCKKTRALFLAVWVLCVFIICFGPTNVLFIHYL 337

Qy   235 PSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSSPFKFYNKLKI 294
                 | | | || : | : || |||||||||||||||| | |: |:|
Db   338 QEANEFLYFAYIL--SACVGSV----------SCCLDPLIYYYASSQCQRYLYSL-L 381

Qy   295 CSLKPKQPGHSKTQ 308
                      ||| | |
Db   382 CCRKVSEPGSSTGQ 395

RESULT 10
I57940
somatostatin receptor 5 - rat
N;Alternate names: somatotropin release-inhibiting factor subtype 2B receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Nov-1999
C;Accession: I57940; I57949; S39244
R;O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.
Mol. Pharmacol. 42, 939-946, 1992
A;Title: Molecular cloning and expression of a pituitary somatostatin receptor with pr

A;Reference number: I57940; MUID:93125499; PMID:1362243
A;Accession: I57940
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-363 <OCA1>
A;Cross-references: GB:L04535; NID:g409238; PIDN:AAA17029.1; PID:g409239
R;O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.
A;Title: Molecular cloning and expression of a pituitary somatostatin receptor with pref
Mol. Pharmacol. 44, 1278, 1993
A;Reference number: I57949; MUID:94088493; PMID:8264565
A;Accession: I57949
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 341-363 <OCA2>
A;Cross-references: GB:S67370; NID:g455947; PIDN:AAB29371.1; PID:g455948
A;Experimental source: pituitary
R;Penetta, R.; Greenwood, M.; Patel, Y.C.
A;Description: Correction of the nucleotide and amino acid sequence of the rat somatosta
submitted to the EMBL Data Library, August 1993
A;Reference number: S39244
A;Accession: S39244
A;Molecule type: mRNA
A;Residues: 309-363 <PEN>
A;Cross-references: EMBL:X74828; NID:g433911; PIDN:CAA52825.1; PID:g433912
C;Genetics:
A;Gene: SSTR5
C;Superfamily: vertebrate rhodopsin

Query Match   16.7%;   Score 309.5;   DB 2;   Length 363;
Best Local Similarity   29.6%;   Pred. No. 5.5e-21;
Matches   96;   Conservative   53;   Mismatches   148;   Indels   27;   Gaps   9;

Qy   17 VMPPLLIVAFVLGALGNGVALCGFCFIMKTWKPSTVYLFNLAVADFLMICLPFRTDYYL 76
               || ::| : |:| ||| ||||||||| :|: ||:
Db   39 LVPVLYLVLVCTVGLSGNTLVIYVLRHAKMKTVNVYILNLAVADVLFMLGLPFLATQNA 98

Qy   77 RRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVAADRFKVVHPHHAVNTISTRVAAGIV 136
               ||| || | || :: : ||||
Db   99 VVSYWPPGSFLCRLVMTLDGINQFTSIFCLMVMSVDRYLAVVHPLAVAIRRWRPRVAKMAS 158

Qy   137 CTLWALVILGTVYLLLENHLCVQETAVSCESFIMESANGWHDIMFQ----LRFFMPLGII 192
               |:| |:| : |:| |: || :|
Db   159 AAVWVFSLLMSLPLLV--FADVQEGWGTCNLSWPEPVGLWGAAPITYTSVLGFFGPLLVI 216

Qy   193 LFCSFKIVWSLR--RRQQLARQARMK-KATRFIMVVAIVFITCYLPSVSARLYFL-WTV 247
               |||| |: | | | ||: :|: || :|: | :|: ||
Db   217 CLCYLLIVVKVKAAGMRVGSSRRRRSEPKVTRMVVVVLVFVGCWLPFFIVNIVNLAFTL 276

Qy   248 PSSACDPSVHGALHITLSPTYMNSMLDPLVYYFSSPSFPKFYNLKIC------SLKP 299
               || : |:| :| : | ::|:|:| : :|:
Db   277 PB--EPTSAGLYFFVVLSYANSCANPILYGFLSDNFRQSFRKV-LCLRRGYGMEDADA 332

Qy   300 KQPGHSKTCRPEEMPISNLGRRSC 323
               | | | || :|| ||
Db   333 IEPRPDKSGRPQ---ATLPTRSC 352

RESULT 11
C41795
somatostatin receptor 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 31-Dec-1993 #text_change 24-Nov-1999
C;Accession: C41795
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A;Title: Cloning and functional characterization of a family of human and mouse somatost
A;Reference number: A41795; MUID:92108031; PMID:1346068
A;Accession: C41795
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-391 <YAM>
A;Cross-references: GB:M81831; NID:g201058; PIDN:AAA58255.1; PID:g201059
C;Superfamily: vertebrate rhodopsin

C;Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein

Query Match   16.6%;   Score 307;   DB 2;   Length 391;
Best Local Similarity   25.7%;   Pred. No. 1e-20;
Matches   85;   Conservative   62;   Mismatches   130;   Indels   54;   Gaps   9;

Qy   4 GSCCRIEGD----------------------TISQVMPPLLIVAFV-----LGALGNGVALCGF 40
               |:|:| :|:|:
Db   24 GACSRGPSGGAADGMBEPGRNASQNGTLSEGQGSAILISFIYSVVCLVGLCGNSMVIYVI 83

Qy   41 CFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLRRRHWARGDIPCRVGLFTLAMNRA 100
               | || |||||||| ||||||| | | | || :| || ||| |:
Db   84 LRYAKMKTATNIYILNLAIADELLMLSVPFLVTSTL-LRHWPFGALLCRLVLSVDAVNMF 142

Qy   101 GSIVFLTVAADRYFKVVHPHHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCV-Q 159
               ||::||| :| || : :|| |:| ||:|| |||||||
Db   143 TSIYCLVLSVDRYVAVVHPIKAARYRRPTVAKVVNLGVWVLSLLVILPIVVFSRTAANS 202

Qy   160 ETAVSCESFIMESANGWHD----IMFQLSFFMPLGIIIFC---------SFKIVWSLR 204
               :: || :   :||:   ::|:|| ::
Db   203 DGTVACNMLMPEPAQRWLVGFVLYTFLMGFLLPVGAICLCYVLIIAKMRMVALKAGWQQR 262

Qy   205 RRQQLARQARMKKATRFIMVVAIVFITCYLPSVSARLYFIWTVPSSACDPSVHGALHITL 264
               : : | | :| || |: :
Db   263 KRSE------RKITLMVMVWVMVFVICWMPFYVVQLVNVFAEQDDAT-----VSQLSV 309

Qy   265 SFTYMNSMLDPLVYYFSSPSFPKFYNKLKIC 295
               | |:| ||:|:   : : :| :
Db   310 ILGTANSCANPILYGFLSDNFKRSFQRI-LC 339

RESULT 12
A41795
somatostatin receptor 1 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Nov-1999
C;Accession: A41795
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A;Title: Cloning and functional characterization of a family of human and mouse somatos
A;Reference number: A41795; MUID:92108031; PMID:1346068
A;Accession: A41795
A;Molecule type: DNA
A;Residues: 1-391 <YAM>
A;Cross-references: GB:M81829; NID:g307433; PIDN:AAA58247.1; PID:g307434
A;Cross-references: GB:M81829; NID:g307434; PIDN:AAA58247.1; PID:g307434
A;Note: sequence extracted from NCBI backbone (NCBIN:74767, NCBIP:74768)
C;Genetics:
A;Gene: GDB:SSTR1
A;Cross-references: GDB:134185; OMIM:182451
A;Map position: 14q13-14q13
A;Introns: #status absent
C;Superfamily: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; ph
F;58-84/Domain: transmembrane #status predicted <TM1>
F;95-120/Domain: transmembrane #status predicted <TM2>
F;132-153/Domain: transmembrane #status predicted <TM3>
F;173-195/Domain: transmembrane #status predicted <TM4>
F;220-250/Domain: transmembrane #status predicted <TM5>
F;269-296/Domain: transmembrane #status predicted <TM6>
F;302-326/Domain: transmembrane #status predicted <TM7>
F;4,44,48,381/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;130-208/Disulfide bonds: #status predicted
F;172/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predi
F;265/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predi
F;339/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match   16.5%;   Score 305.5;   DB 2;   Length 391;
Best Local Similarity   26.3%;   Pred. No. 1.4e-20;
Matches   81;   Conservative   62;   Mismatches   128;   Indels   37;   Gaps   8;

Qy   10 EGDTISQVMPPLLIVAFV-----LGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFL 63
               ||:|:|:   :| ||| ||| ||||||||:|:|||:||||
Db   47 QNGTLSEGQGSAILISFIYSVVCLVGLCGNSMVIYVILRYAKMKTATNIYILNLAIADEL 106

RESULT 13
A39297
somatostatin receptor - rat
N;Alternate names: probable G-protein-coupled receptor; SRIF receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 24-Nov-1999
C;Accession: A39297, A45102, S20088
R;Meyerhof, W.; Paust, H.J.; Schoenrock, C.; Richter, D.
DNA Cell Biol. 10, 689-694, 1991
A;Title: Cloning of a cDNA encoding a novel putative G-protein-coupled receptor expresse
A;Reference number: A39297; MUID:92096119; PMID:1661599
A;Accession: A39297
A;Molecule type: mRNA
A;Residues: 1-391 <MEY>
A;Cross-references: GB:X62314; GB:X61630; NID:g56309; PIDN:CAA44193.1; PID:g56310
A;Experimental source: brain
A;Note: it is uncertain whether Met-1 is the initiator or whether translation is initiat
R;Li, X.J.; Forte, M.; North, R.A.; Ross, C.A.; Snyder, S.H.
J. Biol. Chem. 267, 21307-21312, 1992
A;Title: Cloning and expression of a rat somatostatin receptor enriched in brain.
A;Reference number: A45102; MUID:93016064; PMID:1400442
A;Accession: A45102
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-391 <LII>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIP:116692)
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; receptor; transmembrane protein

Query Match   16.5%;   Score 305.5;   DB 2;   Length 391;
Best Local Similarity   26.3%;   Pred. No. 1.4e-20;
Matches   81;   Conservative   62;   Mismatches 128;   Indels 37;   Gaps   8;

```
Qy    10  EGDTISQVMPPLLIVAFV-----LGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFL    63
Db    47  QNGTLSREGQGSAILISFIYSVVCLVGLCGNSMVIYVILRYAKMKTATNIYILNLAIADEL  106

Qy    64  LMICLPFRTDYILRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHA   123
Db   107  LMLSVPFLVTSTL-LRHWPFGALLCRIVLSVDAVNMFTSIYCLTVLSVDRVVAVVHPIKA   165

Qy   124  VNTISTRVAAGIVCTLWALVILGTVYLLLENHLCV-QETAVSCESFIMESANGWHD----  178
Db   166  ARYRRPTVAKVVNLGVWVLSLVILPIVVFSRTAANSDGTVACNMLMPEPAQRWLVGFVL    225

Qy   179  IMFQLEFFMPLGIILFC--------SFKIVWSLRRRQQLARQARMKKATRPIMVVAI      227
Db   226  YTFLMGFLLPVGAICLCYVLIIAKMRMVALKAGWQQRKRSE-------RKITLMVMWVM    278

Qy   228  VFITCYLPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPK   287
Db   279  VFVICWMPFYVVQLVNVFAEQDDAT-----VSQLSVILGYANSCANPILYGFLSDNFKR    332

Qy   288  FYNKLKIC    295
Db   333  SFQRI-LC    339
```

RESULT 14
S15403
angiotensin II receptor type 1 - bovine
C;Species: Bos primigenius indicus x Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S15403
R;Sasaki, K.; Yamano, Y.; Bardhan, S.; Iwai, N.; Murray, J.J.; Hasegawa, M.; Matsuda,
Nature 351, 230-233, 1991
A;Title: Cloning and expression of a complementary DNA encoding a bovine adrenal angio
A;Reference number: S15403; MUID:91251900; PMID:2041569
A;Accession: S15403
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-359 <SAS>
A;Cross-references: GB:X62294; NID:g43; PIDN:CAA44182.1; PID:g44
C;Superfamily: vertebrate rhodopsin

Query Match   16.4%;   Score 304.5;   DB 2;   Length 359;
Best Local Similarity   24.9%;   Pred. No. 1.6e-20;
Matches   81;   Conservative   68;   Mismatches 143;   Indels 33;   Gaps   8;

```
Qy    17  VMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYIL   76
Db    30  MIPTLYSIIFVWGIFGNSLVVIVIYFYMKLKTVASVFLLNLALADLCFLLTLPLWAVYTA   89

Qy    77  RRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIV  136
Db    90  MEYRWPFGNYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHPMKS-RLRRTMLVAKVT  148

Qy   137  C-TLWALVILGTVYLLL-ENHLCVQETAVSCESFIMESANGWHDIMFQLE-----FFMPL  189
Db   149  CIIWILAGLASLPTIIHRNVFFIENTNITVCAFHYESQNSTLPVGLGLTKNILGFLPF    208

Qy   190  GIILFCSFKIVWSLRRRQQLARQARMKKATRPIMVAIV--FITCYLP----SVSARLYF   243
Db   209  -LILITSYTLIWKTLKKAYBIQKNKPRKDDIFKIIAIVLFFFFSWVPHQIFTFMDVLIQ   267

Qy   244  LWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSPKFY-----------      289
Db   268  LGLIRDCKIEDIVDTAMPIITICLAYPNNCLNPLFYGFLGKKFKKYFLQLLKYIPPKAKSH 327

Qy   290  ---NKLKICSLKPKQPGHSKTQRP   310
Db   328  SNLSTKMSTLSYRPSENGNSSTKKP  352
```

RESULT 15
B41795
somatostatin receptor 2 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Nov-1999
C;Accession: B41795
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A;Title: Cloning and functional characterization of a family of human and mouse somato
A;Reference number: A41795; MUID:92108031; PMID:1346068
A;Accession: B41795
A;Molecule type: DNA
A;Residues: 1-369 <YAM>
A;Cross-references: GB:M81830; NID:g307435; PIDN:AAA58248.1; PID:g307436
A;Note: sequence extracted from NCBI backbone (NCBIN:74769, NCBIP:74770)
C;Genetics:
A;Gene: GDB:SSTR2
A;Cross-references: GDB:134186; CMIM:182452
A;Map position: 17q24-17q24

A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; pho
F;44-69/Domain: transmembrane #status predicted <TM1>
F;80-105/Domain: transmembrane #status predicted <TM2>
F;117-138/Domain: transmembrane #status predicted <TM3>
F;158-180/Domain: transmembrane #status predicted <TM4>
F;205-235/Domain: transmembrane #status predicted <TM5>
F;254-281/Domain: transmembrane #status predicted <TM6>
F;288-315/Domain: transmembrane #status predicted <TM7>
F;9,22,29,32,351/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;115-193/Disulfide bonds: #status predicted
F;250/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predic
F;328/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match   16.3%; Score 302.5; DB 2; Length 369;
Best Local Similarity   29.8%; Pred. No. 2.5e-20;
Matches   90; Conservative   50; Mismatches 135; Indels   27; Gaps   9;

Qy   13 TISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVLFNLAVADFLLMICLPFRT 72
             :|.|| :   :||| || :| ||||:||:||:|| :||:|
Db   41 TSNAVLTFIYFVVCIIGLCGNTLVIYVILRYAKMKTITNIYILNLAIADELFMLGLPFLA 100

Qy   73 DYYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVA 132
             | || ||| :|| |||: || :   :| ||:|| | :||:
Db   101 -MQVALVHWPFGKAICRVVMTVDGINQFTSIPCLTVMSIDRYLAVVHPIKSAKWRRPRTA 159

Qy   133 AGIVCTLWA---LVIIG-TVYLLLENHLCVQETAVSCESFIMESANGWHD----IMFQLE 184
             || : | ::| | :| :| || :   ::
Db   160 KMITMAVWGVSLLVILPIMIYAGLRSN---QWGRSSCTINWPGESGAWYTGFIIYTPILG 216

Qy   185 FFMPLGIILFCSFKIVWSLR---RRQQLARQARMKKATRPIWVWAIVFITCYLPSVSAR 240
             :||| | :: ||| | ||||: |: ||
Db   217 FLVPLTIICLCCYLPIIIKVKSSGIRVGSSKRKKSEKKVTRMVSIVVAVFIFCWLP----- 271

Qy   241 LYFLWTVPSSAC---DPSVHGALHITLSPTYMNSMLDPLVYFSSPSFPKFYNKLKICS 296
             :|||:|: : |: |:: ||:: |: :   : :|
Db   272 -FYIFNVSSVSMAISPTPALKGMFDFVVLJYANSCANPIIYAFLSDNFKKKSFQNV-LCL 329

Qy   297 LK 298
             :|
Db   330 VK 331

Search completed: June 30, 2004, 17:23:53
Job time : 23 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: July 3, 2004, 03:49:46 ; Search time 4415 Seconds
                                     (without alignments)
                                     10190.276 Million cell updates/sec

Title: US-10-076-260-1
Perfect score: 1038
Sequence: 1 atgtacaacgggtcgtgctg........cccacattgttgagtggcac 1038

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
         1: gb_ba:*
         2: gb_htg:*
         3: gb_in:*
         4: gb_om:*
         5: gb_ov:*
         6: gb_pat:*
         7: gb_ph:*
         8: gb_pl:*
         9: gb_pr:*
        10: gb_ro:*
        11: gb_sts:*
        12: gb_sy:*
        13: gb_un:*
        14: gb_vi:*
        15: em_ba:*
        16: em_fun:*
        17: em_hum:*
        18: em_in:*
        19: em_mu:*
        20: em_om:*
        21: em_or:*
        22: em_ov:*
        23: em_pat:*
        24: em_ph:*
        25: em_pl:*
        26: em_ro:*
        27: em_sts:*
        28: em_un:*
        29: em_vi:*
        30: em_htg_hum:*
        31: em_htg_inv:*
        32: em_htg_other:*
        33: em_htg_mus:*
        34: em_htg_pln:*
        35: em_htg_rod:*
        36: em_htg_mam:*
        37: em_htg_vrt:*
        38: em_sy:*
        39: em_htgo_hum:*
        40: em_htgo_mus:*
        41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1038	100.0	1038	6	AX592617	AX592617 Sequence
2	1038	100.0	1038	6	BD140826	BD140826 Novel G p
3	1038	100.0	1041	6	AX147834	AX147834 Sequence
4	1038	100.0	1041	6	AX148182	AX148182 Sequence
5	1038	100.0	1041	6	AX299705	AX299705 Sequence
6	1038	100.0	1041	6	AX375434	AX375434 Sequence
7	1038	100.0	1041	6	AX521883	AX521883 Sequence
8	1038	100.0	1041	6	AX675169	AX675169 Sequence
9	1038	100.0	1041	6	BD144279	BD144279 Novel G-p
10	1038	100.0	1041	9	AB083631	AB083631 Homo sapi
11	1038	100.0	1041	9	AF411110	AF411110 Homo sapi
12	1038	100.0	1050	6	AX338371	AX338371 Sequence
13	1038	100.0	1050	6	AX338373	AX338373 Sequence
14	1038	100.0	1083	6	AX305131	AX305131 Sequence
15	1038	100.0	1372	6	AX549382	AX549382 Sequence
16	1038	100.0	1441	9	AB065866	AB065866 Homo sapi
17	1038	100.0	1730	6	AX277635	AX277635 Sequence
18	1038	100.0	2331	6	AX299707	AX299707 Sequence
19	1038	100.0	2345	9	AF385432	AF385432 Homo sapi
20	1038	100.0	3612	9	AF385431	AF385431 Homo sapi
21 c	1038	100.0	204793	9	AC026333	AC026333 Homo sapi
22	1036.4	99.8	1041	6	AX395171	AX395171 Sequence
23	1036.4	99.8	1194	6	AX395169	AX395169 Sequence
24	1036.4	99.8	2748	9	AK122659	AK122659 Homo sapi
25	1034.8	99.7	1104	6	AX338374	AX338374 Sequence
26	888	85.5	888	6	AX147766	AX147766 Sequence
27	888	85.5	888	6	AX521815	AX521815 Sequence
28	724.4	69.8	1668	6	AX592621	AX592621 Sequence
29	722.8	69.6	3251	6	AX592619	AX592619 Sequence
30	722.8	69.6	256368	2	AC122753	AC122753 Mus muscu
31 c	721.2	69.5	238844	2	AC097683	AC097683 Rattus no
32	578.4	55.7	185956	2	AC129582	AC129582 Mus muscu
33	402.8	38.8	1254	6	AX704524	AX704524 Sequence
34	372.4	35.9	1174	6	AX384661	AX384661 Sequence
35	372.4	35.9	1295	9	AY148884	AY148884 Homo sapi
36	372.4	35.9	1361	9	BD133032	BD133032 HM74 rece
37	369.2	35.6	1092	6	AX148194	AX148194 Sequence
38	369.2	35.6	1092	9	AB083632	AB083632 Homo sapi
39	369.2	35.6	1492	6	AX646475	AX646475 Sequence
40	369.2	35.6	1492	9	AB065876	AB065876 Homo sapi
41	369.2	35.6	2065	9	BC063461	BC063461 Homo sapi
42	367.6	35.4	1092	6	AX704522	AX704522 Sequence
43	367.6	35.4	1164	9	AR308634	AR308634 Sequence
44	367.6	35.4	1564	9	AB065865	AB065865 Homo sapi
45	367.6	35.4	2049	9	BC038955	BC038955 Homo sapi

ALIGNMENTS

RESULT 1
AX592617
LOCUS   AX592617   1038 bp   DNA   linear   PAT 28-JAN-2003
DEFINITION   Sequence 1 from Patent WO02083736.
ACCESSION   AX592617
VERSION   AX592617.1 GI:28144739
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elliott,S.G., Rogers,N. and Busse,L.A.
  TITLE   G-protein coupled receptor molecules and uses thereof
  JOURNAL   Patent: WO 02083736-A 1 24-OCT-2002;

```
FEATURES          Amgen, Inc. (US)
     source       Location/Qualifiers
                  1..1038
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"
     CDS          1..>1038
                  /note="unnamed protein product"
                  /codon_start=1
                  /protein_id="CAD62055.1"
                  /db_xref="GI:28144740"
                  /db_xref="REMTREMBL:CAD62055"
                  /translation="MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHM
                  KTWKPSTVYLFNLAVADFLMKICLPFRTDYLRRRHWAFGDIPCRVGLFTLAMNRAGS
                  IVELTVVAADRYFKVVHPHHAVNTISTRVAAGIVCTLWALVILGTVTLLLRENHLCVQE
                  TAVSCESFIMESANGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKA
                  TRFIMVAIVFITCFLPSVSARLYFLWTVFPSSACDFSVHGALHITLSFTYMNSMLDPL
                  VYYFSSSFPDKFYNKLKICSLKPKQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQS
                  DGQWDPHIVEWH"
ORIGIN

Query Match            100.0%;   Score 1038;   DB 6;   Length 1038;
Best Local Similarity  100.0%;   Pred. No. 1.3e-254;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG 60

Qy     61 CTGCTCATTGTGGCCTTTGTCGTGGCGCACTAGGCAATGGCAATGGGGTCGCCCTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCTCATTGTGGCCTTTGTCGTGGCGCACTAGGCAATGGCAATGGGGTCGCCCTGTGGTTTC 120

Qy    121 TGCTTCCACATGAAGACCTGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTTCCACATGAAGACCTGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCT 180

Qy    181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGAGACTATTACCTCAGACGTAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATTTCCTCCTTATGATCTGCCTGCCTTTCGGACAGAGACTATTACCTCAGACGTAGACAC 240

Qy    241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGGCTCTTCACGTTGCCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGGCTCTTCACGTTGCCCATGAACAGGGCC 300

Qy    301 GGGAGCATCGTGTTCCTTACGGTGTGGGCTGCGACCAGGTATTTCAAAGTGGTCCACCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATCGTGTTCCTTACGGTGTGGGCTGCGACCAGGTATTTCAAAGTGGTCCACCCC 360

Qy    361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGGCGCTGGCATCGTCGCATCGTCGCACCCGTGTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGGCGCTGGCATCGTCGCATCGTCGCACCCGTGTGG 420

Qy    421 GCCCCTGGTCATCTCGGAACAGTGTATCTTTGCTGGAGAACCATCTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCCTGGTCATCTCGGAACAGTGTATCTTTGCTGGAGAACCATCTGCGTGCAAGAG 480

Qy    481 ACGGGCCGTCTCCGTGTGAGAGCTTCATCATGAGAGCTCGGCCAATGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGGCCGTCTCCGTGTGAGAGCTTCATCATGAGAGCTCGGCCAATGGCTGGCATGACATCATG 540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600

Qy    601 TGGAGCCTGAGGCGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGGCGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACCCGG 660

Qy    661 TTCATCATGGTGGTGGCCAATTGGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGGTGGCCAATTGGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy    721 CTCTATTTCCTCTGGACGGTGCCCTGAGTGCTGCGATCCCTCGTCCATGGGCCCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGGACGGTGCCCTGAGTGCTGCGATCCCTCGTCCATGGGCCCTG 780

Qy    781 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGACATCGTGGAGTGTATTATTTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGACATCGTGGAGTGTATTATTTT 840

Qy    841 TCAAGCCCCTCCTTCCCAAATTCTACACAAGTCAACAAATCTGCAGTCTGAAACCCAAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGCCCCTCCTTCCCAAATTCTACACAAGTCAACAAATCTGCAGTCTGAAACCCAAG 900

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTCGAACCTCGGTCGC 960

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTGGCAATGGGGATCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTGGCAATGGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038
```

RESULT 2
BD140826
LOCUS   BD140826   1038 bp   DNA   linear   PAT 18-SEP-2002
DEFINITION Novel G protein-coupled receptor protein and its DNA.
ACCESSION   BD140826
VERSION   BD140826.1 GI:23235771
KEYWORDS   WO 0202767-A/1.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1038)
  AUTHORS   Shibata,S., Horikoshi,K., Taniyama,Y., Shintani,Y. and Miyajima,N.
  TITLE   Novel G protein-coupled receptor protein and its DNA
  JOURNAL   Patent: WO 0202767-A 1 10-JAN-2002;
            TAKEDA CHEMICAL INDUSTRIES LTD, SACHIO SHIBATA, KENICHI HORIKOSHI,
            YOSHIO TANIYAMA, YASUSHI SHINTANI, NOBUYUKI MIYAJIMA
  OS   Homo sapiens (human)
  PN   WO 0202767-A/1
  PD   10-JAN-2002
  PF   02-JUL-2001 WO 2001JP005711
  PR   04-JUL-2000 JP 00P 206860, 31-JUL-2000 JP 00P 235274 PI
            SACHIO SHIBATA, KENICHI HORIKOSHI, YOSHIO TANIYAMA, YASUSHI PI
            SHINTANI,
  PI   NOBUYUKI MIYAJIMA
  PC   C12N15/09, C12N15/12, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C07K14/ PC
            705,
  PC   C07K16/28, C12P21/02, C12Q1/02, C12Q1/68, A61K45/00, A61P29/00, PC
            A61P3/04,
  PC   A61P43/00, G01N33/15, G01N33/50
  CC   Novel G protein-coupled receptor protein and its DNA FH   Key
FT   source   1..1038
FT   /organism='Homo sapiens (human)'
FEATURES   Location/Qualifiers
     source   1..1038
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%;   Score 1038;   DB 6;   Length 1038;
Best Local Similarity 100.0%;   Pred. No. 1.3e-254;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG 60
```

```
LOCUS           AX147834                          1041 bp    DNA        linear   PAT 08-JUN-2001
DEFINITION      Sequence 79 from Patent WO0136473.
ACCESSION       AX147834
VERSION         AX147834.1  GI:1436838
KEYWORDS        .
SOURCE          Homo sapiens (human)
  ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1
  AUTHORS       Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
                Slighton,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
                Seilitz,T. and Huff,R.M.
  TITLE         Novel g protein-coupled receptors
  JOURNAL       Patent: WO 013647J-A 79 25-MAY-2001;
                PHARMACIA & UPJOHN COMPANY (US)
FEATURES        Location/Qualifiers
     source     1..1041
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN

Query Match        100.0%;    Score 1038;   DB 6;   Length 1041;
Best Local Similarity   100.0%;    Pred. No. 1.3e-254;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps     0;

Qy         1 ATGTACAACGGGTCGTGTGCCGCATGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGTACAACGGGTCGTGTGCCGCATGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60

Qy        61 CTGCTCATTGTGCCTTTGTGCTGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CTGCTCATTGTGCCTTTGTGCTGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120

Qy       121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTACCTTTFCAATTTGGCCGTGGCT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTACCTTTFCAATTTGGCCGTGGCT 180

Qy       181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACACGACTATTACCTCAGACGTAGACAC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACACGACTATTACCTCAGACGTAGACAC 240

Qy       241 TGGGCTTTGGGACATTCCTGCGAGTGGGGTCTTCACGTTGGCCATGAACAGGGGCC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TGGGCTTTGGGACATTCCTGCGAGTGGGGTCTTCACGTTGGCCATGAACAGGGGCC 300

Qy       301 GGGAGCATCGTGTTCCTTACGGTGGTGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GGGAGCATCGTGTTCCTTACGGTGGTGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360

Qy       361 CACCACGCCGGTGAACACTATCTCGCGTGCAAGAG 420
             ||||||||||||||||||||||||||||||||||||
Db       361 CACCACGCCGGTGAACACTATCTCGCGTGCAAGAG 420

Qy       421 GCCCTGGTCATCCTGGGAACACTATCTCACCCGGTGGCCGGCTGCGACCCCTGTGG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GCCCTGGTCATCCTGGGAACACTATCTCACCCGGTGGCCGGCTGCGACCCCTGTGG 480

Qy       481 ACGGCCGTCTCCGTGTCCTCGTGTGAGAGCTTCATCATGGCTGGCCAATGGCATCATG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 ACGGCCGTCTCCGTGTCCTCGTGTGAGAGCTTCATCATGGCTGGCCAATGGCATCATG 540

Qy       541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCCTTCAAGATTGTT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCCTTCAAGATTGTT 600

Qy       601 TGGAGCCTGAGGCGGAGGCCAGCAGCTGGCTCGAGTGAAGAGGCGACCCGG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TTCATCATGGTGGTGCAATTGTGTTCATCACATGTCACCTGCCCCAGCGTGTCTGCCAGA 720
```

Db   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACACAGGGCC 300

Qy   301 GGGAGCATGTGTTCCTTACGGTGGCTGGCTGGCTGGCAGGTATTTCAAAGTGGTCCACCCC 360
Db   301 GGGAGCATGTGTTCCTTACGGTGGCTGGCTGGCTGGCAGGTATTTCAAAGTGGTCCACCCC 360

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGTGGCGGCTGGCGCTGCGACCCTGTGG 420
Db   361 CACCACGCGGTGAACACTATCTCCACCCGGTGGCGGCTGGCGCTGCGACCCTGTGG 420

Qy   421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTCGGAGAACCATCTCTGCCTGCAAGAG 480
Db   421 GCCCTGGTCATCCTGGAACAGTGTATCTTTGCTCGGAGAACCATCTCTGCCTGCAAGAG 480

Qy   481 ACGGCCGTCTCCGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCCATGACATCATG 540
Db   481 ACGGCCGTCTCCGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCCATGACATCATG 540

Qy   541 TTCCAGCTGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGTCCTTCAAGATTGTT 600
Db   541 TTCCAGCTGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCAGCAGCTCCGAGCAGGCTCGGATGAAGAAGGGCGACCCGG 660
Db   601 TGGAGCCTGAGGCGGAGGCAGCAGCTCCGAGCAGGCTCGGATGAAGAAGGGCGACCCGG 660

Qy   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
Db   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy   721 CTCTATTTCCTCTGGACGGTGCCCTGCGATCCCTGCGAGTGCCTGTCCATGGGCCCTG 780
Db   721 CTCTATTTCCTCTGGACGGTGCCCTGCGATCCCTGCGAGTGCCTGTCCATGGGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACCTACACGAACAGACATGCTGGATCCCCGGTGTATTATTT 840
Db   781 CACATAACCCTCAGCTTCACCTACACGAACAGACATGCTGGATCCCCGGTGTATTATTT 840

Qy   841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAGCTCACAAATCTGCAGTCTGAAACCCAAG 900
Db   841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAGCTCACAAATCTGCAGTCTGAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
Db   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCCAATGGGATCCC 1020
Db   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 5
AX299705
LOCUS   AX299705   1041 bp   DNA   linear   PAT 26-NOV-2001
DEFINITION   Sequence 1 from Patent WO0173029.
ACCESSION   AX299705
VERSION   AX299705.1   GI:17129250
KEYWORDS   .
SOURCE   Homo sapiens (human)
   ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
   AUTHORS   Ye,J.C., Cravchik,A.C., di Francesco,V.C. and Beasley,E.M.
   TITLE   Isolated human g-protein coupled receptors, nucleic acid molecules
            encoding human gpcr proteins, and uses thereof
   JOURNAL   Patent: WO 0173029-A 1 04-OCT-2001;
            PB Corporation (NY) (US)
FEATURES   Location/Qualifiers

Db   661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy   721 CTCTATTTCCTCTGGACGGTGCCCTGCGATCCCTGCGAGTGCCTGTCCATGGGCCCTG 780
Db   721 CTCTATTTCCTCTGGACGGTGCCCTGCGATCCCTGCGAGTGCCTGTCCATGGGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACCTACACGAACAGACATGCTGGATCCCCGGTGTATTATTT 840
Db   781 CACATAACCCTCAGCTTCACCTACACGAACAGACATGCTGGATCCCCGGTGTATTATTT 840

Qy   841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAGCTCACAAATCTGCAGTCTGAAACCCAAG 900
Db   841 TCAAGCCCCTCCTTCCCAAATTCTACACAACAGCTCACAAATCTGCAGTCTGAAACCCAAG 900

Qy   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
Db   901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCCAATGGGATCCC 1020
Db   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 4
AX148182
LOCUS   AX148182   1041 bp   DNA   linear   PAT 08-JUN-2001
DEFINITION   Sequence 23 from Patent WO0136471.
ACCESSION   AX148182
VERSION   AX148182.1   GI:14347084
KEYWORDS   .
SOURCE   Homo sapiens (human)
   ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
   AUTHORS   Chen,R., Dang,H.T. and Lowitz,K.P.
   TITLE   Endogenous and non-endogenous versions of human g protein-coupled
            receptors
   JOURNAL   Patent: WO 0136471-A 23 25-MAY-2001;
            Arena Pharmaceuticals, Inc. (US)
FEATURES   Location/Qualifiers
   source   1..1041
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN

Qy   1021 CACATTGTTGAGTGGCAC 1038
Db   1021 CACATTGTTGAGTGGCAC 1038

Query Match   100.0%;   Score 1038;   DB 6;   Length 1041;
Best Local Similarity   100.0%;   Pred. No. 1.3e-254;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCGCGCATCGCCATCTCCCAGGTGGATGCGCGCCG 60
Db   1 ATGTACAACGGGTCGTGCGCGCATCGCCATCTCCCAGGTGGATGCGCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGGTTTC 120
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCCACTGTTTACCTTTCAATTTGGGCCGTGGCT 180
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCCACTGTTTACCTTTCAATTTGGGCCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACGACTATTACCTCAGACGACTGGACAC 240
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACGACTATTACCTCAGACGACTGGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACACAGGGCC 300

source   1..1041
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%;   Score 1038;   DB 6;   Length 1041;
Best Local Similarity   100.0%;   Pred. No. 1.3e-254;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCCACTAGGCAATGGGCTCGCCCTGTGTGGTTTC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCCACTAGGCAATGGGCTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTTACCTTTCAATTTGGCGTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTTACCTTTCAATTTGGCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGCGCTCTTCACGTTGGCCATGAACAGGGCC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGCGCTCTTCACGTTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360

Qy   361 CACCACGCGGGTGAACACTATCTCCACCCGGGTGGCGTGGCATCGTCGCACCCTGTGG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGGGTGAACACTATCTCCACCCGGGTGGCGTGGCATCGTCGCACCCTGTGG 420

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGAGTCGGCCAATGGCTGGCATGACATCATG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGAGTCGGCCAATGGCTGGCATGACATCATG 540

Qy   541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTGTCCTTCAAGAATTGTT 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTGTCCTTCAAGAATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCAGCAGCAGTCGGCCAGACAGCTCGGATGAAGAAGGCGACCCGG 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCAGCAGCAGTCGGCCAGACAGCTCGGATGAAGAAGGCGACCCGG 660

Qy   661 TTCATCATGGTGGTGGCCAATGTGTTCATCACATCACATGCCAGCGTGCTCTGCTAGA 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCATCATGGTGGTGGCCAATGTGTTCATCACATCACATGCCAGCGTGCTCTGCTAGA 720

Qy   721 CTCTATTTCCTCGGAGCGGTGGCCCTCGAGTGCCTGCCATCCCTCGTCCATGGGGCCTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTCTATTTCCTCGGAGCGGTGGCCCTCGAGTGCCTGCCATCCCTCGTCCATGGGGCCTG 780

Qy   781 TGCTTCCACATGAAGACCTGGAAGCCCACTACATGAACACCGGTGTATTATTTT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TGCTTCCACATGAAGACCTGGAAGCCCACTACATGAACACCGGTGTATTATTTT 840

Qy   841 CACATAACCCTCAGCTTCACCTACAGAAAGGCCGGAAGATGCCAATTTCGAACCTCGGTCGC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CACATAACCCTCAGCTTCACCTACAGAAAGGCCGGAAGATGCCAATTTCGAACCTCGGTCGC 900

Qy   901 CAGCCAGGAGACACTCAAAAACACAAAAGGCCGGAAGATGCCGAATTTCGAACCTCGGTCGC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGCCAGGAGACACTCAAAAACACAAAAGGCCGGAAGATGCCGAATTTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTCGATGGGCAATGGGATCCC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTCTCGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
           ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 6
AX375434
LOCUS   AX375434   1041 bp   DNA   linear   PAT 01-MAR-2002
DEFINITION Sequence 11 from Patent WO0198330.
ACCESSION   AX375434
VERSION   AX375434.1 GI:19170025
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Lannoy,V., Brezillon,S., Detheux,M., Parmentier,M. and Govarts,C.
TITLE   A recombinant cell line expressing gpcrxll as a functional receptor
            validated by angiopeptin and useful for screening of agonists and
            antagonists
JOURNAL   Patent: WO 0198330-A 11 27-DEC-2001;
            Euroscreen S.A. (BE)
FEATURES   Location/Qualifiers
     source   1..1041
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS   1..1041
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD26778.1"
                     /db_xref="GI:19170026"
                     /db_xref="REMTREMBL:CAD26778"
                     /translation="MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGCFFM
                     KTWKPSTVYLPNLAVADFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLFTLAMNRAGS
                     IVFLTVVAADRYFKVVHPHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQE
                     TAVSCESFIMESANGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKA
                     TRFIMVVAIVFITCYLPSVSARIYFLWTVPSSACDPSVRGALHITLSFTYMNSMLDPL
                     VYFSSSFPKFYNKLKICSLKZKQPGHSKTQPEEMPISNLGRRSCISVANSFCSQS
                     DGQWDPHIVEWH"
ORIGIN

Query Match   100.0%;   Score 1038;   DB 6;   Length 1041;
Best Local Similarity   100.0%;   Pred. No. 1.3e-254;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCCACTAGGCAATGGGCTCGCCCTGTGTGGTTTC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCCACTAGGCAATGGGCTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTTACCTTTCAATTTGGCCGTGGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTTACCTTTCAATTTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGCGCTCTTCACGTTGGCCATGAACAGGGCC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGCGCTCTTCACGTTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTTCCACCCC 360

ORIGIN

Query Match   100.0%; Score 1038; DB 6; Length 1041;
Best Local Similarity 100.0%;
Matches 1038; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ATGTACAACGGGTCGTGCTGCGCATGCGACGGGGACCACCATCTCCAGGTGATGCGCGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTACAACGGGTCGTGCTGCGCATGCGACGGGGACCACCATCTCCAGGTGATGCGCGCCG 60

Qy     61 CTGCTCATTGTGCCTTGTGCGGCGGCCACTAGGCGACTAGGCGACTGTTACCTTTCGTGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCTCATTGTGCCTTGTGCGGCGGCCACTAGGCGACTAGGCGACTGTTACCTTTCGTGTTTC 120

Qy    121 TGCTTCCAATGAAGACCTGGAAGCCAGACACTGTTTACCTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTTCCAATGAAGACCTGGAAGCCAGACACTGTTTACCTTTCAATTGGCCGTGGCT 180

Qy    181 GATTTCCTCTTATGATCTGCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATTTCCTCTTATGATCTGCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy    241 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC 300

Qy    301 GGGAGCATCGTGTTCCTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGAGCATCGTGTTCCTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC 360

Qy    361 CACCACGCGGTGAACACTATCTCCACCCGGTGGGCGTCTGGCATCGTCTCGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGCGGTGAACACTATCTCCACCCGGTGGGCGTCTGGCATCGTCTCGTG 420

Qy    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy    481 ACGGCCGTCTCGTGTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCGTGTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG 540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTATTATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTATTATTGTT 600

Qy    601 TGGAGCCTGAGGCGGAGGCAGCAGCCAGACACAGGCTCGGATGAAGAAGGCGACCCGG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGGCGGAGGCAGCAGCCAGACACAGGCTCGGATGAAGAAGGCGACCCGG 660

Qy    661 TTCATCATGGTGGCAATTGTGTTCATCACATGCTTGCGATCCCTCGTCCTGCCCAGCGTGTCTGCTAGA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCATGGTGGCAATTGTGTTCATCACATGCTTGCGATCCCTCGTCCTGCCCAGCGTGTCTGCTAGA 720

Qy    721 CTCTATTTCCTCTGGACGGTGCCCTGCGAGTGCCTTGCGGATCCCGTTCCATGGGGCCCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTATTTCCTCTGGACGGTGCCCTGCGAGTGCCTTGCGGATCCCGTTCCATGGGGCCCTG 780

Qy    781 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGATCCCCTGGTGTATTATTTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGATCCCCTGGTGTATTATTTT 840

Qy    841 TCAAGGCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAAGGCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG 900

Qy    901 CAGCCAGACACTCAAAAACAACAAAGGCCGGAAGATGCCAATTTCGAACCTCGGTCGC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CAGCCAGACACTCAAAAACAACAAAGGCCGGAAGATGCCAATTTCGAACCTCGGTCGC 960

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTTCGATGGGCAATGGGATCCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCAAAGCCAGTTCGATGGGCAATGGGATCCC 1020

Db    301 GGGAGCATCGTGTTCCTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGGTCCACCCC 360

Qy    361 CACCACGCGGTGAACACTATCTCCACCCGGTGGGCGTCTGGCATCGTCTCGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACCACGCGGTGAACACTATCTCCACCCGGTGGGCGTCTGGCATCGTCTCGTG 420

Qy    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG 480

Qy    481 ACGGCCGTCTCGTGTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACGGCCGTCTCGTGTCCTGTGAGAGCTTCATCATGGCTGGCATGACATCATG 540

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTATTATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCAGCTGGAGTTCTTTATGCCCCTGGATCCCTGGTGTATTATTGTT 600

Qy    601 TGGAGCCTGAGGCGGAGGCAGCAGCCAGACACAGGCTCGGATGAAGAAGGCGACCCGG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGGAGCCTGAGGCGGAGGCAGCAGCCAGACACAGGCTCGGATGAAGAAGGCGACCCGG 660

Qy    901 CAGCCAGGACACTCAAAAAACAACAAAGGCCGGAAGATGTGGCAAATAGTTTCCAAAGCTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGAT 1020
Db    901 CAGCCAGGACACTCAAAAAACAACAAAGGCCGGAAGATGTGGCAAATAGTTTCCAAAGCTGTGGCAAATAGTTTCCAAAGCCAGTCGATGGGCAATGGGAT 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038
```

RESULT 7
AX521883

LOCUS   AX521883   1041 bp   DNA   linear   PAT   24-OCT-2002
DEFINITION Sequence 79 from Patent WO02064789.
ACCESSION   AX521883
VERSION   AX521883.1 GI:24410790
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lind,P., Parodi,L.A., Vogeli,G. and Wood,L.S.
  TITLE   G protein-coupled receptor
  JOURNAL   Patent: WO 02064789-A 79 22-AUG-2002;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES   Location/Qualifiers
     source   1..1041
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

```
Qy   601  TGGAGCCTGAGGCGGAGGCAGCAGCGCTGGCCAGACAGGCTTCGGATGAAGAGAGGCGACCCGG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TGGAGCCTGAGGCGGAGGCAGCAGCGCTGGCCAGACAGGCTTCGGATGAAGAGAGGCGACCCGG  660

Qy   661  TTCATCATGGTGGTGGCAATTGTGTTCATCACATGTCTACCTGCCCAGCCTGTGCTGCTAGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TTCATCATGGTGGTGGCAATTGTGTTCATCACATGTCTACCTGCCCAGCCTGTGCTGCTAGA  720

Qy   721  CTCTATTCCTCTGGACGGTGCCCTCGAGTGCCTGGCATCCTCTGTCCATGGGGCCTG  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CTCTATTCCTCTGGACGGTGCCCTCGAGTGCCTGGCATCCTCTGTCCATGGGGCCCTG  780

Qy   781  CACATAACCCTCAGCTTCACCTACATGACAGCATGCTGGATCCCCTGGTGTATTATTTT  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CACATAACCCCTCAGCTTCACCTACATGACAGCATGCTGGATCCCCTGGTGTATTATTTT  840

Qy   841  TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TCAAGCCCCTCCTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG  900

Qy   901  CAGCCAGGAGCACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTGAACCTCGGTCGC  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CAGCCAGGAGCACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTGAACCTCGGTCGC  960

Qy   961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGGCAATGGGATCCC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGGCAATGGGATCCC  1020

Qy  1021  CACATTGTTGAGTGGCAC  1038
          ||||||||||||||||||
Db  1021  CACATTGTTGAGTGGCAC  1038
```

RESULT 8
AX675169
LOCUS   AX675169   1041 bp   DNA   linear   PAT 27-MAR-2003
DEFINITION   Sequence 5 from Patent WO03002604.
ACCESSION   AX675169
VERSION   AX675169.1 GI:29933412
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bhatia,U., Jones,C.E., Bouhelal,R., Seuwen,K. and Tenaillon,L.
  TITLE   Novel G protein-coupled receptors and dna sequences thereof
  JOURNAL   Patent: WO 03002604-A 5 09-JAN-2003;
             Novartis AG (CH)
FEATURES   Location/Qualifiers
     source   1..1041
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%; Score 1038; DB 6; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.3e-254;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1  ATGTACAACGGGTCGTGCTGCGCCATCGAGAGGGGGACACCATCTCCCAGGTGATGCCGCCG  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGTACAACGGGTCGTGCTGCGCCATCGAGAGGGGGACACCATCTCCCAGGTGATGCCGCCG  60

Qy    61  CTGCTCATTGTGGCCTTGTCTGCTGGGCCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CTGCTCATTGTGGCCTTGTCTGCTGGGCCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC  120

Qy   121  TGCTTCACCATGAAGAGACCTGGAAGCCCAGCACTGTTACCTTTTACCTTTTGGCCGTGGCT  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  TGCTTCCACCATGAAGAGACCTGGAAGCCCAGCACTGTTACCTTTTACCTTTTCAATTGGCCCGTGGCT  180

Qy   181  GATTTCCTCCTTATGATCTGCCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GATTCCCCTTATGATCTGCCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240

Qy   241  TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGTCTTCACGTTGGCCATGAACAGAGGGC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGTCTTCACGTTGGCCATGAACAGAGGGC  300

Qy   301  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCCGGACAGTGTATTCAAAGTGGTCCACCCC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCCGGACAGTGTATTCAAAGTGGTCCACCCC  360

Qy   361  CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGCGGTGGCATCGTCGCACCCTGTGG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGCGGTGGCATCGTCGCACCCTGTGG  420

Qy   421  GCCCTCGGTCATCCTGGGAACAGTGTATCTTTGCTGGGAGAACCATCTCGCGTGCAAGAG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGTGAGAACCATCTCGCGTGCAAGAG  480

Qy   481  ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCATGACATCATG  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ACGGCCGTCTCCTGTGAGAGCTTCATCATGAGTCGGCCAATGGCTGGCATGACATCATG  540

Qy   541  TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TTCCAGCTGGAGTTCTTTATGCCCCTGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
```

RESULT 9
BD144279
LOCUS   BD144279   1041 bp   DNA   linear   PAT 17-JAN-2003
DEFINITION   Novel G-protein coupled receptors.
ACCESSION   BD144279
VERSION   BD144279.1 GI:27850037
KEYWORDS   JP 2002112793-A/4.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1041)
  AUTHORS   Haga,T., Takeda,S. and Miyake,N.
  TITLE   Novel G-protein coupled receptors
  JOURNAL   Patent: JP 2002112793-A 4 16-APR-2002;
             JAPAN SCIENCE AND TECHNOLOGY CORP
COMMENT   OS   Homo sapiens (human)
             PN   JP 2002112793-A/4
             PD   16-APR-2002
             PF   09-FEB-2001 JP 2001034434
             PI   TATSUYA HAGA, SHIGEKI TAKEDA, NARIKI MIYAKE
             PC
             C12N15/09,A01K67/027,A61K38/00,A61K39/395,A61K45/00, PC
             A61K48/00,
             PC   A61P43/00,C07K14/705,C07K16/28,C07K19/00,C12N1/15,C12N1/19, PC
             C12N1/21,
             PC   C12N5//10,C12Q1/02,C12Q1/68,G01N33//15,G01N33/50,G01N33/53, PC
             G01N33/566//
             PC   C12P21/08,C12N15//00,A61K37/02,C12N5/00
             CC   Novel G-protein coupled receptors
             FH   Key   Location/Qualifiers
             FT   CDS   (1)..(1041).
FEATURES   Location/Qualifiers
     source   1..1041
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%; Score 1038; DB 6; Length 1041;

us-10-076-260-1.rge

Wed Jul 7 12:07:41 2004

Best Local Similarity 100.0%; Pred. No. 1.3e-254;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTCGTGCGTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCGCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCGTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCGCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTACTGTTTACCTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACAGGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACAGGGCC 300

Qy   301 GGGAGCATGCGTGTTCCTTACGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGAGCATGCGTGTTCCTTACGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 360

Qy   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCACCCTGTGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGCACCCTGTGG 420

Qy   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCGCGTGCAAGAG 480

Qy   481 ACGGCCGTCTCCTGTGGAGAGCTTCATCATGGAGTCGGCCAATGCTGGCATGACATCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACGGCCGTCTCCTGTGGAGAGCTTCATCATGGAGTCGGCCAATGCTGGCATGACATCATG 540

Qy   541 TTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCCAGCTGGAGTTCTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT 600

Qy   601 TGGAGCCTGAGGCGGAGGCCAGCTGCCAGACAGGCTCGGATGAAGAGGCGACCCCGG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TGGAGCCTGAGGCGGAGGCCAGCTGCCAGACAGGCTCGGATGAAGAGGCGACCCCGG 660

Qy   661 TTTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy   721 CTCTATTTCCTCTGGACGGTGCCCCTGCAGTGCTTGTCCATGGGGCCCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTCTATTTCCTCTGGACGGTGCCCCTGCAGTGCTTGTCCATGGGGCCCTG 780

Qy   781 CACATAACCCTCAGCTTCACCTACATGAACTACATGCTGAAAGCCCCGGTGTATTATTTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CACATAACCCTCAGCTTCACCTACATGAACTACATGCTGAAAGCCCCGGTGTATTATTTT 840

Qy   841 TCAAGCCCCTCCTTTCCAAATTCTACACAACAAGCTCAGTCTGAAACCCAAG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCAAGCCCCTCCTTTCCAAATTCTACACAACAAGCTCAGTCTGAAACCCAAG 900

Qy   901 CAGGCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CAGGCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC 960

Qy   961 AGGAGTTGCATCAGTGTGCAAATAGTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AGGAGTTGCATCAGTGTGCAAATAGTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020

Qy   1021 CACATTGTTGAGTGGCAC 1038
          ||||||||||||||||||
Db   1021 CACATTGTTGAGTGGCAC 1038

Db   1021 CACATTGTTGAGTGGCAC 1038

RESULT 10

LOCUS   AB083631   1041 bp   DNA   linear   PRI 24-MAY-2002
DEFINITION   Homo sapiens GPCR gene for putative G-protein coupled receptor,
                complete CDS, clone:hGPCR49.
ACCESSION   AB083631
VERSION   AB083631.1 GI:20152325
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Takeda,S., Kadowaki,S., Haga,T., Takaesu,H. and Mitaku,S.
  TITLE   Identification of G protein-coupled receptor genes from the human
                genome sequence
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1041)
  AUTHORS   Takeda,S., Kadowaki,S., Haga,T., Takaesu,H. and Mitaku,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-APR-2002) Shigeki Takeda, Gunma University, Faculty of
                Department of Biological and Chemical, Engineering, Gunma University,
                Engineering, 1-5-1, Kiryu, Gunma 376-8515, Japan
                (E-mail:stakeda@bce.gunma-u.ac.jp, Tel:+81-277-30-1434,
                Fax:+81-277-30-1434)
FEATURES   Location/Qualifiers
     source   1..1041
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="hGPCR49"
     gene   1..1041
                     /gene="GPCR"
     CDS   1..1041
                     /gene="GPCR"
                     /note="predicted with SOSUI analysis"
                     /codon_start=1
                     /product="putative G-protein coupled receptor"
                     /protein_id="BAB89344.1"
                     /db_xref="GI:20152326"
                     /translation="MYNGSCRIRGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHM
                     KTWKPSTVYLFNLAVADFLIMICLPFRTDYLRRHMAFGDIPCRVGLFTLAMNRAGS
                     IVPLTVVAADRYPFKVVHPHAVNTISTRVAAGIVCTLWALVLILGTVYLLLENHLCVQE
                     TAVSCBSFIMESANGWHDIMFQLEFFMPLGILLPCSFKIVWSLRRRQQLARQARMKKA
                     TRFIMVVAIVFITCYLPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPL
                     VYYFSSPSFPKFYNKLKICSLKPQPGHSKTQRPEEMPISNLGRRSCISIVANSFQSQS
                     DGQWDPHIVEWH"
ORIGIN

Db   1021 CACATTGTTGAGTGGCAC 1038

Query Match   100.0%; Score 1038; DB 9; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.3e-254;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTACAACGGGTCGTGCGTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCGCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACAACGGGTCGTGCGTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCGCGCG 60

Qy   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTC 120

Qy   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCTTCCACATGAAGACCTGGAAGCCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT 180

Qy   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC 240

Qy   241 TGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTGGCCATGAACAGGGCC 300

PUBMED   11574155
REFERENCE   2 (bases 1 to 1041)
AUTHORS   Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,
   Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.
TITLE   Direct Submission
JOURNAL   Submitted (17-AUG-2001) Department of Pharmacology, University of
   Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada
FEATURES   Location/Qualifiers
source   1..1041
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="12"
   /map="12q"
gene   <1..>1041
   /gene="GPR81"
mRNA   <1..>1041
   /gene="GPR81"
   /product="G protein-coupled receptor"
CDS   1..1041
   /gene="GPR81"
   /codon_start=1
   /product="G protein-coupled receptor"
   /protein_id="AAL26481.1"
   /db_xref="GI:16656326"
   /translation="MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFEM
   KTWKPSTVLFLNLAVADFLLMICLPFRTDYYLRRHWAFGDIPCRVGLFFLAMNRAGS
   IVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQE
   TAVSCESFIMESANGWHDIMFQLRFFMPLGIILPCSFKIVWSLRRRQQARQARMKKA
   TRFIMVVAIVFITCYLPSVSARLYFLW'VPSSACDPSVHGALHITLSFYMNSMLDPL
   VYYFSSPSFPKFYNKLKICSLKPKQPGHSKTQPEEMPISNLGRRSCISVANSFQSQS
   DGQWDPHIVEWH"

ORIGIN

Query Match   100.0%;   Score 1038;   DB 9;   Length 1041;
Best Local Similarity   100.0%;   Pred. No. 1.3e-254;
Matches 1038;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy    1  ATGTACAACGGGTCGTGCTGCCGCATGGAGGGGGACACCATCTCCCAGGTGATGCCGCCG  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGTACAACGGGTCGTGCTGCCGCATGGAGGGGGACACCATCTCCCAGGTGATGCCGCCG  60

Qy   61  CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTGGCCCTGTGCGGTTTC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTGGCCCTGTGCGGTTTC  120

Qy  121  TGCTTCCAATGAAGACCTGGAAGCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  180

Qy  181  GATTTCCTCTTATGATCGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GATTTCCTCTTATGATCGCCTTTTCGGACAGACTATTACCTCAGACGTAGACAC  240

Qy  241  TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300

Qy  301  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGTCCACCCC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGTCCACCCC  360

Qy  361  CACCACGCCGGTGAACACTATCTCCACCCGGGTGGCGGCCTCGCATCGCTGCACCCGTGG  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CACCACGCCGGTGAACACTATCTCCACCCGGGTGGCGGCCTCGCATCGCTGCACCCGTGG  420

Qy  421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480

Qy  481  ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGGCTGGCATGACATCATG  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ACGGCCGTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGGCTGGCATGACATCATG  540
```

```
Db  241  TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGTCCACCCC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTCAAAGTGTCCACCCC  360

Qy  361  CACCACGCCGGTGAACACTATCTCCACCCGGGTGGCGGCCTCGCATCGCTGCACCCGTGG  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CACCACGCCGGTGAACACTATCTCCACCCGGGTGGCGGCCTCGCATCGCTGCACCCGTGG  420

Qy  421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480

Qy  481  TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCCTGCCAGAGCTCGGATGAAGAAGGCGACCCGG  600

Db  541  TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTGTCTCTTCAAGATTGTT  600

Qy  601  TGGAGCCTGAGGCGGAGGCGGAGGCGCAGCAGTGGCCAGACAGCTCGGATGAAGAAGGCGACCCGG  660

Db  601  TGGAGCCTGAGGCGGAGGCGGAGGCGCAGCAGTGGCCAGACAGCTCGGATGAAGAAGGCGACCCGG  660

Qy  661  TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA  720

Db  661  TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGA  720

Qy  721  CTCTATTTCCTCTGGACGGTGCCCTGCGAGTGCCTGCGATCCCCTCGTCCATGGGGCCCTG  780

Db  721  CTCTATTTCCTCTGGACGGTGCCCTGCGAGTGCCTGCGATCCCCTCGTCCATGGGGCCCTG  780

Qy  781  CACATAACCCTGAGCTTCACCTCAGCTTCACCTACACATGAACACAGCATGCTGGATCCCCTGGTGTATTATTTT  840

Db  781  CACATAACCCTGAGCTTCACCTCAGCTTCACCTACACATGAACACAGCATGCTGGATCCCCTGGTGTATTATTTT  840

Qy  841  TCAAGCCCCTCCTTCTTCCCAAATTCTACACAACAGCTCAAAATCTGCAGTCGAAACCCAAG  900

Db  841  TCAAGCCCCTCCTTCTTCCCAAATTCTACACAACAGCTCAAAATCTCAGTCGAAACCCAAG  900

Qy  901  CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTTCGGTCGC  960

Db  901  CAGCCAGGACAGTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTTCGGTCGC  960

Qy  961  AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTGCAATGGCGAATGGGATCCC  1020

Db  961  AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTGCAATGGGCAATGGGATCCC  1020

Qy 1021  CACATTGTTGAGTGGCAC  1038

Db 1021  CACATTGTTGAGTGGCAC  1038
```

RESULT 11
AF411110
LOCUS   AF411110   1041 bp   DNA   linear   PRI 01-NOV-2001
DEFINITION   Homo sapiens G protein-coupled receptor (GPR81) gene, complete cds.
ACCESSION   AF411110
VERSION   AF411110.1   GI:16566325
KEYWORDS   Homo sapiens (human)
SOURCE   Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1041)
AUTHORS   Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,
   Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.
TITLE   Discovery and mapping of ten novel G protein-coupled receptor genes
JOURNAL   Gene 275 (1), 83-91 (2001)
MEDLINE   21458557

```
Qy     541 TTCCAGCTGGAGTTCTTTATGCCCCTCCGGCATCATCTTATTTTGTCTCCTTCAAGATTGTT 600
               ||||||||||||||||||| ||||||||||||||| ||||||||| |||||||||||||||
Db     541 TTCCAGCTGGAGTTCTTTATGCCCCTCCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 600

Qy     601 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGACGTGGCCATGGGGTCGG 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGACGTGGCCATGGGGTCGG 660

Qy     661 TTCATCATGGTGGTGGCAATTGTGTTCATCACAGCTACCTGCCCAGCGTGTCTGCTAGA 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TTCATCATGGTGGTGGCAATTGTGTTCATCACAGCTACCTGCCCAGCGTGTCTGCTAGA 720

Qy     721 CTCTATTTCCTCTGGACGGTGCCCTCCGAGTGCGTGGCCTCGGATCCCTGTCCATGGGCCCTG 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTCTATTTCCTCTGGACGGTGCCCTCCGAGTGCGTGGCCTCGGATCCCTGTCCATGGGGCCCTG 780

Qy     781 CACATAACCCTCCTTCCTACACAACAGCATGCTGATCCCCGTGGTATTATTT 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CACATAACCCTCCTTCCTACACAACAGCATGCTGATCCCCGTGGTATTATTT 840

Qy     841 TCAAGCCCCTCCTTCTCTACAAATTCTACAACAAGCTCAAAATTGCAGTCTGAAACCCAAG 900
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCAAGCCCCTCCTTCTCTACAAATTCTACAACAAGCTCAAAATTGCAGTCTGAAACCCAAG 900

Cy     901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCTGGATCCCCTGGTGTATTATTT 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCTGGATCCCCTGGTGTATTATTT 960

Qy     961 AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTCTGAATGGGCAATGGGATCCC 1020
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AGGAGTTGCATCAGTGTGGCAAATAGTTCCAAAGCCAGTCTGAATGGGCAATGGGATCCC 1020

Qy    1021 CACATTGTTGAGTGGCAC 1038
               ||||||||||||||||||
Db    1021 CACATTGTTGAGTGGCAC 1038
```

RESULT 12
AX338371
LOCUS   AX338371   1050 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION Sequence 1 from Patent WO0174904.
ACCESSION   AX338371
VERSION   AX338371.1 GI:18128869
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Majumder,K., Vernet,C.A., Casman,S.J., Wolenc,A.R., Spaderna,S.K.,
            Padigaru,M., Mishmu,V.S., Tchernev,V.T., Spytek,K.A., Li,L.,
            Baumgartner,J.C. and Gusev,V.Y.
  TITLE   Novel proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0174904-A 1 11-OCT-2001;
            Curagen Corporation (US)
FEATURES   Location/Qualifiers
     source   1..1050
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%; Score 1038; DB 6; Length 1050;
Best Local Similarity 100.0%;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6 ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGTGATGCCGCCG 65

Qy      61 CTGCTCATTGTGGCCTTTGTGCTGTCGGCGCCACTAGGCAATGGGGTCGCCCGTGTGGTTTC 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Page 10 alignments (top section):

```
Db      66 CTGCTCATTGTGGCCTTTGTGCTGTCGGCCGCACTAGGCAATGGGGTCGCCCGTGTGGTTTC 125

Qy     121 TGCTTCCACATGAAGAGACCTGGAAGCCCAGCCCAGCACTGTTGTTACTT TTGGCCGTGGCT 180
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 TGCTTCCACATGAAGAGACCTGGAAGCCCAGCCCAGCACTGTTTACCTTTTGGCCGTGGCT 185

Qy     181 GATTTCCTCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 240
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 GATTTCCTCTTATGATCTGCCTGCCTTTCGGACAGACTATTACCTCAGACGTAGACAC 245

Qy     241 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGTCTTCACGTTGGCCATGAACAGGGCC 300
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     246 TGGGCTTTTGGGACATTCCCTGCCGAGTGGGGTCTTCACGTTGGCCATGAACAGGGCC 305

Qy     301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCCC 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC 365

Qy     361 CACCAGCGCGTGAACACTATCTCCACCCGGGTGCGGCGCTGGCATCGTCTCGCGTGCAAGAG 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 CACCAGCGCGTGAACACTATCTCCACCCGGGTGCGGCGCTGGCATCGTCTCGCGTGCAAGAG 425

Qy     421 GCCCTGGTCATCCTGGGACAGTGTATCTTTGTGGAGAACCATCTCGCGTGCATGACATCATG 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 GCCCTGGTCATCCTGGGACAGTGTATCTTTGTGGAGAACCATCTCGCGTGCATGACATCATG 485

Cy     481 ACGGCCGTCTCCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     486 ACGGCCGTCTCCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATG 545

Qy     541 TTCCAGCTGGAGTTCTTTATGCCCCTCCGGCATCATCTTATTTTGTCTCCTTCAAGATTGTT 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     546 TTCCAGCTGGAGTTCTTTATGCCCCTCCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT 605

Qy     601 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGACGTGGCCATGGGGTCGG 660
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     606 TGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGACGTGGCCATGGGGTCGG 665

Qy     661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGTTCATCACATGTGCCACTGCCCTGCTAGA 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     666 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGTTCATCACATGTGCCACTGCCCTGCTAGA 725

Qy     721 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCGTGCCCTCGGATCCCTGTCCATGGGGCCCTG 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     726 CTCTATTTCCTCTGGACGGTGCCCTCGAGTGCGTGCCCTCGGATCCCTGTCCATGGGGCCCTG 785

Qy     781 CACATAACCCTCAGCTTCCACTACACAACAAGCTCAAATCTGAGATCCCCTGGTGTATTATTT 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     786 CACATAACCCTCAGCTTCCACTACACAACAAGCTCAAATCTGAGATCCCCTGGTGTATTATTT 845

Qy     841 TCAAGCCCCTCCTTCCCAAATTCCCAAATTCTACAACAAGCTCAAAATCTGAGAGTCTGAAACCCAAG 900
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     846 TCAAGCCCCTCCTTCCCAAATTCCCAAATTCTACAACAAGCTCAAAATCTGAGAGTCTGAAACCCAAG 905

Qy     901 CAGCCAGGAGACACTCAGCTTCACCTACATGAACAGGCCGGAAGAGATGCCGGAACCTCGGTCGC 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     906 CAGCCAGGAGACACTCAGCTTCACCTACATGAACAGGCCGGAAGAGATGCCGGAACCTCGGTCGC 965

Qy     961 AGGAGTTGCATCAGTAGTTCCAAAGCCAGTCTGAATGGGCAATGGGATCCC 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     966 AGGAGTTGCATCAGTAGTTCCAAAGCCAGTCTGAATGGGCAATGGGATCCC 1025

Qy    1021 CACATTGTTGAGTGGCAC 1038
               ||||||||||||||||||
Db    1026 CACATTGTTGAGTGGCAC 1043
```

RESULT 13
AX338373
LOCUS   AX338373   1050 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION Sequence 3 from Patent WO0174904.
ACCESSION   AX338373
VERSION   AX338373.1 GI:18128870

```
Db   726  CTCTATTTCCTCTGGACGGTGCCTCGAGTGCTGCGGATCCCTCGTCCATGGGGCCCTG   785
                                                                       
Qy   781  CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGGTGTATTATTT   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786  CACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCTCGGTGTATTATTT   845

Qy   841  TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846  TCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG  905

Qy   901  CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906  CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC  965

Qy   961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCCAATGGGATCCC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCCAATGGGACCCC 1025

Qy  1021  CACATTGTTGAGTGGCAC  1038
          ||||||||||||||||||
Db  1026  CACATTGTTGAGTGGCAC  1043
```

RESULT 14

```
AX305131
LOCUS       AX305131                1083 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 11 from Patent WO0187937.
ACCESSION   AX305131
VERSION     AX305131.1  GI:17644766
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Patterson,C., Lu,D.A., Thornton,M., Lu,Y., Tribouley,C.M.,
            Graul,R., Khan,F.A., Gandhi,A.R., Walia,N.K., Nguyen,D.B., Yue,H.,
            Hafalia,A., Elliott,V.S., Lal,P., Reddy,R., Kallick,D.A., Tang,T.Y.
            and Au-Young,J.
  TITLE     G-protein coupled receptors
  JOURNAL   Patent: WO 0187937-A 11 22-NOV-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1083
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID NO: 7474846CB1"
ORIGIN
```

```
Query Match          100.0%;  Score 1038;  DB 6;  Length 1083;
Best Local Similarity 100.0%;  Pred. No. 1.3e-254;
Matches 1038;  Conservative   0;  Mismatches  0;  Indels   0;  Gaps   0;

Qy     1  ATGTACAACGGGTCGTGCTGCCGCATCGAGAGGGGACACCATCTCCCAGGTGATGCGGCCG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6  ATGTACAACGGGTCGTGCTGCCGCATCGAGAGGGGACACCATCTCCCAGGTGATGCGGCCG   65

Qy    61  CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCGTGTGGTTTC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66  CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCGTGTGGTTTC  125

Qy   121  TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126  TGCTTCCACATGAAGACCTGGAAGCCAGCACTGTTTACCTTTCAATTGGCCGTGGCT  185

Qy   181  GATTTCCTCTCTATGATCTGCCTGCCTTTGGACAGACTATTACCTTCAGACGTAGACAC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186  GATTTCCTCTCTATGATCTGCCTGCCTTTGGACAGACTATTACCTTCAGACGTAGACAC  245

Qy   241  TGGGCTTTTGGGGACATTCCCTGCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246  TGGGCTTTTGGGGACATTCCCTGCGAGTGGGGCTCTTCACGTTGGCCATGAACAGGGCC  305

Qy   301  GGGAGCATCGTGTTCCTTACGTGGTGGCTGCGCGACAGGTATTTCAAAGTGGTCCACCCC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306  GGGAGCATCGTGTTCCTTACGTGGTGGCTGCGCGACAGGTATTTCAAAGTGGTCCACCCC  365

Qy   361  CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGGCATCGTCTGCTGCACCCGTGTGG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366  CACCACGCGGTGAACACTATCTCCACCCGGGTGCGGCTGGCATCGTCTGCTGCACCCGTGTGG  425

Qy   421  GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426  GCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAG  485

Qy   481  ACGGCCGTCTCCGTGAGAGCTTCATCATGGAGTCGGCCAATGGCCAATGACATCATG  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486  ACGGCCGTCTCCGTGAGAGCTTCATCATGGAGTCGGCCAATGGCCAATGACATCATG  545

Qy   541  TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546  TTCCAGCTGGAGTTCTTTATGCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTT  605

Qy   601  TGGAGCCTGAGCGCGGAGGCAGCAGTGGCCAGAGGCTCGGATGAAGAGAGGGGACCCGG  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606  TGGAGCCTGAGCGCGGAGGCAGCAGTGGCCAGAGGCTCGGATGAAGAGAGGGGACCCGG  665

Qy   661  TTCATCATGGTGGTGGCAATTGGCCAATTGTGTTCATCACATGCTACCTGCCCCAGCGC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666  TTCATCATGGTGGTGGCAATTGGCCAATTGTGTTCATCACATGCTACCTGCCCCAGCGC  725

Qy   721  CTCTATTTCCTCTGGACGGTGCCTCGAGTGCTGCGGATCCCTCGTCCATGGGGCCCTG  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726  CTCTATTTCCTCTGGACGGTGCCTCGAGTGCTGCGGATCCCTCGTCCATGGGGCCCTG  785
```

source

1..1372
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match   100.0%;   Score 1038;   DB 6;   Length 1372;
Best Local Similarity   100.0%;   Pred. No. 1.3e-254;
Matches 1038;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ATGTACAACGGGTCGTGCTGCCGCATCGAGCGGGGACACCATCTCCAGTGATGCGGCCG    60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 ATGTACAACGGGTCGTGCTGCCGCATCGAGCGGGGACACCATCTCCAGTGATGCGGCCG   201

Qy     61 CTGCTCATTGTGGGCCTTTGTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGTTTC   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 CTGCTCATTGTGGGCCTTTGTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGTTTC   261

Qy    121 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTACCTTTTCAATTGGCCGGTGGCT   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 TGCTTCCACATGAAGACCTGGAAGCCCAGACACTGTTTACCTTTTCAATTGGCCGGTGGCT   321

Qy    181 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGAGACTATTACCTCAGACGTAGAGACAC   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 GATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGAGACTATTACCTCAGACGTAGAGACAC   381

Qy    241 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGCTCTTCAGTGGCCATGAACAGGGGCC   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 TGGGCTTTTGGGGACATTCCTGCCGAGTGGGCTCTTCAGTGGCCATGAACAGGGGCC   441

Qy    301 GGGAGCATCGTGTTCCTACGGTGGTGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    442 GGGAGCATCGTGTTCCTACGGTGGTGCTGCGGACAGGTATTTCAAAGTGGTCCACCCC   501

Qy    361 CACCACGCGGTGAACATATCTCCACCCCGGGTGGCGGCTCGCCGCATCGTCTGCACCCGTGG   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 CACCACGCGGTGAACATATCTCCACCCCGGGTGGCGGCTCGCCGCATCGTCTGCACCCGTGG   561

Qy    421 GCCCTGGTCATCCTGGAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGTGCGCTGTGG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 GCCCTGGTCATCCTGGAACAGTGTATCTTTTGCTGGAGAACCATCTCGCGTGCGCTGTGG   621

Qy    481 ACGGCCGCTCTCGTGAGAGCTTCATCATGGCTGGCCAATGGCTGGCATGACATCATG   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 ACGGCCGCTCTCGTGAGAGCTTCATCATGGCTGGCCAATGGCTGGCATGACATCATG   681

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATGCTGATGGGCAATGGGGATCCC   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATGCTGATGGGCAATGGGGATCCC   741

Qy    601 TGGAGCCTGAGCGCGGAGGCAGCAGCCTCGGATGAAGAGGCGACCCGG   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 TGGAGCCTGAGCGCGGAGGCAGCAGCCTCGGATGAAGAGGCGACCCGG   801

Qy    661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGTGTTCATCACATGTTCTGCTAGA   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGTGTTCATCACATGTTCTGCTAGA   861

Qy    721 CTCTATTTCCTCTGGACGGTGCCCTGAGTGGCTTCATCGAGTCCCTGTCCATGGGCCCTG   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 CTCTATTTCCTCTGGACGGTGCCCTGAGTGGCTTCATCGAGTCCCTGTCCATGGGCCCTG   921

Qy    781 CACATAACCCTCAGCTTCACCTACATGAACAC   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 CACATAACCCTCAGCTTCACCTACATGAACAC   981

Qy    841 TCAAGCCCCTCCTTTCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 TCAAGCCCCTCCTTTCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG   1041

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGGTCGC   960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGGTCGC   1101

Qy    301 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGGACAGGTATTTCAAAGTGGTCCACCCC   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    343 GGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGGACAGGTATTTCAAAGTGGTCCACCCC   402

Qy    361 CACCACGCGGTGAACACTATCTCCACCCCGGGTGGCGGCTGGCGGGACAGCAGGTATTTTCAAAGTGGTCCACCCTGTGG   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    403 CACCACGCGGTGAACACTATCTCCACCCCGGGTGGCGGCTGGCGGGACAGCAGGTATTTTCAAAGTGGTCCACCCTGTGG   462

Qy    421 GCCCTGGTCATCCTGGAACAGTGTATCTTTTTGCTGGAGAACCATCTCGCGTGCTGCAAGAG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    463 GCCCTGGTCATCCTGGGAACAGTGTATCTTTTTGCTGGAGAACCATCTCGCGTGCTGCAAGAG   522

Qy    481 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGCTGGCCAATGGCTGGCATGACATCATG   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    523 ACGGCCGTCTCCTGTGAGAGCTTCATCATGGCTGGCCAATGGCTGGCATGACATCATG   582

Qy    541 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    583 TTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTGCTCCTTCAAGATTGTT   642

Qy    601 TGGAGCCTGAGGGCGGAGGCAGCGCTGGCCAGACAGGCTCGGACACAGGGCCGACCCGG   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    643 TGGAGCCTGAGGGCGGAGGCAGCGCTGGCCAGACAGGCTCGGACACAGGGCCGACCCGG   702

Qy    661 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTCACCTGCCCAGCGTGTCTGCTAGA   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    703 TTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTCACCTGCCCAGCGTGTCTGCTAGA   762

Qy    721 CTCTATTTCCTCTGGACGGTGCCCTGAGTGGCTGCCTCGGGATCCCTGTCGTTATTATTT   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    763 CTCTATTTCCTCTGGACGGTGCCCTGAGTGGCTGCCTCGGGATCCCTGTCGTTATTATTT   822

Qy    781 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGATCTGGTGTATTATTTT   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    823 CACATAACCCTCAGCTTCACCTACATGAACAGCAGCATGCTGGATCTGGTGTATTATTTT   882

Qy    841 TCAAGCCCCTCCTTTCCCAAATTCTCACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    883 TCAAGCCCCTCCTTTCCCAAATTCTCACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAG   942

Qy    901 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC   960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    943 CAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGAACCTCGGTCGC   1002

Qy    961 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGGATCCC   1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1003 AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGGATCCC   1062

Qy   1021 CACATTGTTGAGTGGCAC   1038
            ||||||||||||||||||
Db   1063 CACATTGTTGAGTGGCAC   1080
```

RESULT 15
AX549382

LOCUS   AX549382   1372 bp   DNA   linear   PAT 26-NOV-2002
DEFINITION   Sequence 667 from Patent WO02061087.
ACCESSION   AX549382
VERSION   AX549382.1   GI:25814001
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Burmer,G.C., Roush,C.L. and Brown,J.P.
TITLE   Antigenic peptides, such as for G protein-coupled receptors
           (GPCRs), antibodies thereto, and systems for identifying such
           antigenic peptides
JOURNAL   Patent: WO 02061087-A 667 08-AUG-2002;
           Lifespan Biosciences, Inc. (US)
FEATURES   Location/Qualifiers

```
Qy    961  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102  AGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCC 1161

Qy   1021  CACATTGTTGAGTGGCAC 1038
           ||||||||||||||||||
Db   1162  CACATTGTTGAGTGGCAC 1179
```

Search completed: July 3, 2004, 07:48:28
Job time : 4423 secs